

1 GCCATGGTGG GGCAGAGGTT GGGGAAGATGG CGTGGCGAGG CTGGGCGCAG
51 AGAGGCTGGG GCTGCGGCCA GGCCTGGGGT GCGTCGGTGG GCGGCCGAG
101 CTGCGAGGAG CTCAGTCCGG TCCTAACCCC GCCGCAGCTC CTCGGACGCA
151 GGTTTAACTT CTTTATTCAA CAAAAATGCG GATTTCAGAAA AGCACCCAGG
201 AAGGTTGAAC CTCGAAGATC AGACCCAGGG ACAAGTGGTG AAGCATACAA
251 GAGAAGTGCT TTGATTCTCT CTGTGGAAGA AACAGTCTTT TATCCTTCTC
301 CCTATCCTAT AAGGAGTCTC ATAAAACCTT TATTTTTTAC TGTGGGTTT
351 ACAGGCTGTG CATTTGGATC AGCTGCTATT TGGCAATATG AATCACTGAA
401 ATCCAGGGTC CAGAGTTATT TTGATGGTAT AAAAGCTGAT TGGTTGGATA
451 GCATAAGACC ACAAAAAGAA GGAGACTTCA GAAAGGAGAT TAACAAGTGG
501 TGGAATAACC TAAGTGATGG CCAGCGGACT GTGACAGGTA TTATAGCTGC
551 AAATGTCCTT GTATTCTGTT TATGGAGAGT ACCTTCTCTG CAGCGGACAA
601 TGATCAGATA TTTACATCG AATCCAGCCT CAAGTGTTAT TTCCAATTTT
651 GTCAGTTACG TGGGTAAAGT TGCCACAGGA AGATATGGAC CATCACTTGG
701 TGCATCTGGT GCCATCATGA CAGTCCTCGC AGCTGTCTGC ACTAAGATCC
751 CAGAAGGGAG GCTTGCCATT ATTTTCCTTC CGATGTTTAC GTTACAGCA
801 GGGAATGCCC TGAAAGCCAT TATCGCCATG GATACAGCAG GAATGATCCT
851 GGGATGGAAG TTTTGTGATC ATGCGGCACA TCTTGGGGGA GCTCTTTTGT
901 GAATATGGTA TGTTACTTAC GGTCATGAAC TGATTTGGAA GAACAGGGAG
951 CCGCTAGTGA AAATCTGGCA TGAAATAAGG ACTAATGGCC CCAAAAAAGG
1001 AGGTGGCTCT AAGTAAACT GGGATTGGAC AGTAGTGGTG CATCTGGTCC
1051 TTGCCGCTG AGAGCCCCAG GAGACATCGG CTAGAGTGAC CATGGCTATG
1101 CTCCCGTCTG GAAGATGCCA GCATCTGGCC TCCCACTGTT TTCAGCTGTG
1151 TCCCCAGTC CGTGTCTTTT TAGAATGTGA ATGATGATAA AGTTGTGAAA
1201 TAAAGGTTTC TATCTAGTTT GTAAAAA AAAAAA AAAAAA (SEQ ID NO:1)

FEATURES:

5'UTR: 1 - 26
Start Codon: 27
Stop Codon: 1014
3'UTR: 1017

Homologous proteins:

gi 11066250 gb AAG28519.1 AF197937_1 (AF197937) presenilins int...	668	0.0
gi 8924134 ref NP_061092.1 hypothetical protein PRO2207 [Homo ...	264	1e-69
gi 7303544 gb AAF58598.1 (AE003824) CG8972 gene product [Droso...	186	4e-46
gi 3219925 sp O14364 YB4J_SCHPO HYPOTHETICAL 33.6 KD PROTEIN C3...	69	1e-10
gi 6321538 ref NP_011615.1 Ygr101wp [Saccharomyces cerevisiae]...	64	3e-09

FIGURE 1A

EST:

gi 10216540 /dataset=dbest /taxon=96...	1203	0.0
gi 10215044 /dataset=dbest /taxon=96...	1203	0.0
gi 10212049 /dataset=dbest /taxon=96...	1172	0.0
gi 10154606 /dataset=dbest /taxon=96...	1160	0.0
gi 9141009 /dataset=dbest /taxon=9606...	1144	0.0
gi 9338606 /dataset=dbest /taxon=960...	1094	0.0
gi 9720819 /dataset=dbest /taxon=960...	1090	0.0
gi 5857747 /dataset=dbest /taxon=9606 ...	1033	0.0
gi 10813749 /dataset=dbest /taxon=960...	1009	0.0

EXPRESSION INFORMATION FOR MODULATORY USE:

gi|10216540 Lung
gi|10215044 Lung small cell carcinoma
gi|10212049 Lung small cell carcinoma
gi|10154606 Ovary adenocarcinoma
gi|9141009 Lung
gi|9338606 Uterus endometrium
gi|9720819 Lymph Burkitt lymphoma
gi|5857747 Colon
gi|10813749 Dendritic cells

Tissue Expression:

Human leukocytes

1 MAWRGWAQRG WCGGQAWGAS VGGRSCEELT AVLTPPQLLG RRFNFFIQQK
51 CGFRKAPRKV EPRRSDPGTS GEAYKRSALI PPVEETVFYP SPYPIRSLLK
101 PLFFTVGFTG CAFGSAAIWQ YESLKSRVQS YFDGIKADWL DSIRPQKEGD
151 FRKEINKWWN NLSGQRTVT GIIAANVLVF CLWRVPSLQR TMIRYFTSNP
201 ASSVISNFVS YVGKVATGRY GPSLGASGAI MTVLAHVCTK IPEGRLAIF
251 LPMFTFTAGN ALKAIAMDT AGMILGWKFF DHAHLGGAL FGIWVYTYGH
301 ELIWKNNREPL VKIWHIIRTN GPKKGKGGSK (SEQ ID NO:2)

FEATURES:

Functional domains and key regions:

Prosite results:

[1] PDOC00001 PS00001 ASN_GLYCOSYLATION

N-glycosylation site

161-164 NLS

[2] PDOC00005 PS00005 PKC_PHOSPHO_SITE

Protein kinase C phosphorylation site

Number of matches: 3

- 1 123-125 SLK
- 2 142-144 SIR
- 3 217-219 TGR

[3] PDOC00006 PS00006 CK2_PHOSPHO_SITE

Casein kinase II phosphorylation site

Number of matches: 3

- 1 25-28 SCEE
- 2 69-72 TSGE
- 3 130-133 SYFD

[4] PDOC00008 PS00008 MYRISTYL

N-myristoylation site

Number of matches: 10

- 1 12-17 GCGQAW
- 2 14-19 GQAWGA
- 3 18-23 GASVGG
- 4 22-27 GGRSCE
- 5 110-115 GCAFGS
- 6 171-176 GIIAAN
- 7 225-230 GASGAI
- 8 228-233 GAIMTV

FIGURE 2A

9 272-277 GMILGW
10 288-293 GALFGI

[5] PDOC00009 PS00009 AMIDATION
Amidation site

39-42 LGRR

Membrane spanning structure and domains:

Helix	Begin	End	Score	Certainty
1	107	127	1.825	Certain
2	173	193	1.069	Certain
3	226	246	1.654	Certain
4	250	270	1.382	Certain
5	288	308	1.123	Certain

BLAST Alignment to Top Hit:

Alignment to top blast hit:

>gi|11066250|gb|AAG28519.1|AF197937_1 (AF197937) presenilins
interacting rhomboid-like protease [Homo sapiens]
Length = 379

Score = 668 bits (1706), Expect = 0.0

Identities = 327/379 (86%), Positives = 328/379 (86%), Gaps = 50/379 (13%)

Frame = +3

Query: 27 MAWRGWAQRGWGCGQAWGASVGGRSCEELTAVLTPPQLLGRRFNFFIQQKCGFRKAPRKV 206
MAWRGWAQRGWGCGQAWGASVGGRSCEELTAVLTPPQLLGRRFNFFIQQKCGFRKAPRKV
Sbjct: 1 MAWRGWAQRGWGCGQAWGASVGGRSCEELTAVLTPPQLLGRRFNFFIQQKCGFRKAPRKV 60

Query: 207 EPRRSDPGTSGEAYKRSALIPPVEETVFYPSYPYIRSLIKPLFFTVGFTGCAFGSAAIWQ 386
EPRRSDPGTSGEAYKRSALIPPVEETVFYPSYPYIRSLIKPLFFTVGFTGCAFGSAAIWQ
Sbjct: 61 EPRRSDPGTSGEAYKRSALIPPVEETVFYPSYPYIRSLIKPLFFTVGFTGCAFGSAAIWQ 120

Query: 387 YESLKSRVQSYFDGIKADWLDSIRPQKEGDFRKEINKWNNLSDGQRTVTGIIAANLVF 566
YESLKSRVQSYFDGIKADWLDSIRPQKEGDFRKEINKWNNLSDGQRTVTGIIAANLVF
Sbjct: 121 YESLKSRVQSYFDGIKADWLDSIRPQKEGDFRKEINKWNNLSDGQRTVTGIIAANLVF 180

Query: 567 CLWRVPSLQRTMIRYFTSNPAS----- 632
CLWRVPSLQRTMIRYFTSNPAS
Sbjct: 181 CLWRVPSLQRTMIRYFTSNPASKVLCSPMLLSTFSHFSLFHMAANMYVLWSFSSSIIVNIL 240

Query: 633 -----SVISNFVSYGKVVATGRYGPSLGASGAIMTVLAAVCTKIPEGRLAIIF 776
VISNFVSYGKVVATGRYGPSLGASGAIMTVLAAVCTKIPEGRLAIIF
Sbjct: 241 GQEQFMAVYLSAGVISNFVSYLGVKVVATGRYGPSLGASGAIMTVLAAVCTKIPEGRLAIIF 300

FIGURE 2B

Query: 777 LPMFTFTAGNALKAIIAMDTAGMILGWKFFDHAHLGGALFGIWYVTYGHელიWKNREPL 956
LPMFTFTAGNALKAIIAMDTAGMILGWKFFDHAHLGGALFGIWYVTYGHელიWKNREPL
Sbjct: 301 LPMFTFTAGNALKAIIAMDTAGMILGWKFFDHAHLGGALFGIWYVTYGHელიWKNREPL 360

Query: 957 VKIWHEIRTNGPKKGGGSK 1013
VKIWHEIRTNGPKKGGGSK
Sbjct: 361 VKIWHEIRTNGPKKGGGSK 379 (SEQ ID NO:4)

Hmmer search results (Pfam):

Scores for sequence family classification (score includes all domains):

Model	Description	Score	E-value	N
PF01694	Rhomboid family	23.3	1.8e-05	1

Parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
PF01694	1/1	201	292 ..	59	147 ..	23.3	1.8e-05

FIGURE 2C

1 CGAGGTTTCT TCATGTTGGT CAGGCTGGTC TCGAACTCCC GACCTCAGGT
51 GATCCGTC CGCCTTTAT CTTTCATTTT TTTTCATGTA TTTTCCTTTA
101 CCACCGCACC CGGCCTTTAT CTTTCATTTT TTTTCATGTA TTTTCCTTTA
151 TTTTAATCAC TTTATCCAGA AACATATCCT CGTCTTGACA GTGCTGTGGT
201 GCCTGTGGTT TCCAGAAGCT GGGTGTGCTG TGTGTCTGTG GTTTGAGGAA
251 GTTGCCCATG GAACTGACAG AGGAAGCAGA GTAGTCGTTG CCATTTTTC
301 GCCTAGTAGG CAGGATCAGG GACCCCATCT TGCTCTCTTT GCCTTGAACC
351 ACAATTAGAA TAAAACACCA AAGCCCTGAC TGATCATGAT CATAGCAATC
401 CGATCTTTAT GATCATGGCC AGACCATTCT CAGGTCGTCT TTACCCTAAG
451 ATATCAATCA CTGGGTATGA CAACCTAGAC CTAAGGGTGC ACTCTGGGTA
501 GTAAAGATGA TTAACCTCTC CAAAGGAATC TAAGGAATCC AGAGCAACAC
551 GAATCACTGC TCTCTTCTA TAGGGTAAAC CTCCAAGAC TCCAGTCCCT
601 GTGAGGAGGC TCTGCCCGCC TGCCCTTCCC AGGGTCCAG GCTCCACATT
651 GGGAGGTGTA CACAGTGCTC TTCGCTCTTC ATTGCCTTGT GTATGATCCC
701 TTTTCCCATC TTTGCATAAA TGCTGTCCCT CTCACCATCT TTAAAAGAGT
751 TCTGGGTAAT TATTTACCAA AGGTGGTATA ATGCTGTCAC AGTCCCTGCT
801 AGTGAGACAT CTGATACAAC TGATGGAATC AGTTCAACAA AATGCAGTAA
851 AATTTTATTT AATGTACTAC GGAGAAAGAA AAAATGCTAC CAGTTATAAG
901 ATGCATCCTG ATTTTCAGATA TTTAAATGGA AAAATGTCT TAAGATCTGT
951 GAAAAATGTA GCTTCTTTC CCACCTCTCA AGTGGGAGAG CAAAAACTGG
1001 ACAGACTAGA AATGCCAGGG GCTAGCTGAG AACCTTACAG AATGAGCAAC
1051 TGCGGAAGCC ACAGGTAACA CCGAGATGTA GATCAGCTGC CAGGGACAAG
1101 ACAAAGAATG TTTTCTAAAG TAAATCCTCT TACCAGTATG TTATTGAAAT
1151 CAGTCCTTAT TGGCATCGAA GAAGGTGAAA GTGCTACTTG CCTGTTGCCT
1201 ACAGAGACTG GAGGAATGAC AAATGTTTAA ATTATTTTAA TTCAACAAGT
1251 AGAGGAATAC CTGCTATGTG AAGGAGTTGT GGCAATTCAT AAAATTAATA
1301 TATTTTTTGA AGTTTGTAGT TTTCAATAAT AATTTCTTAT CTAAAATGTA
1351 ACAAGTTAAT TATATTATCG AATAAACCTC AATTTCTAG TACTAACAAC
1401 ATCAACACTT ACAGAAAAAG GAAAGTCACT CAACTCCCAC ATGTAAACAG
1451 ACTTTAGAAG CAGTTGCAGA GGTTTTCTAA ATTATCCCTG AATTCCTATC
1501 ACATGACTAT TTTTCTCAGA CATGTTGACC TTCACCTACA CAGATGACTC
1551 ACATATGTTT CCATAAGCTG GCAGTAAGTT TAAGAAGCAT ACCATGCCCT
1601 GAGGAAAAAG AAGTAATGTT AGCTCTTCTA CTCTTGGCCA AAGAACCTAA
1651 TTCTGTATAT TACTTCTGTC TTTGGTTTGG CTATTATAGA CAATAAATTA
1701 TTGATCTGAT TATAATTGAG AAAAGTAAGC TCTTCTAAAG AAGTAAAATA
1751 TGGATCTAGG GAAAGGAAGT TAGCTCCAG AGCATTTACA ATTTCCAGG
1801 AATTCTGTGA CTTTACCAAC CCTAGGCAGT GCTGATACTT TAAAAGCATT
1851 CATTTCACTT GCTTTTTTTT GGCTCACCCC CTATCCCCCA GGTATACAGT
1901 ACTCTTACAT AATTGTGGAA GAATCTTACA AGGGGGTAAT GTAGATCAGA
1951 CTTTCTGCT TTCATTTTAA ACCTCCCTAA ATTATAAATA TTTATTTTGT
2001 AGGTATTATA GCTGCAAATG TCCTTGATTT CTGTTTATGG AGAGTACCTT
2051 CTCTGCAGCG GACAATGATC AGATATTTCA CATCGAATCC AGCCTCAAGT
2101 AAGTCTAACT TGTGTGAATT TATTTTAAGG TAGAAATAAT ATGAAAGAAA
2151 TATGCTTTAG TTAATGGAAG TGCTGTAAAA AAGACGAATT ACCTATCAAT
2201 AGCTACAAGC AAAATGCAGA GGATAGGCTG TAAGCTCCTT CACTGAGGAC
2251 AGGGACCTCA CCTCTCTTTT TCTTTTCTT TGTTTTTTTT GAGACGGAGT

FIGURE 3A

2301 CTTCCCTCTGT TGCCCAGGCT GGAGTGCAGT GGTGCAGTCT TAGCTCACTA
2351 CAACCTCCAC CTCCCAGGTT CAAGTGATTCT CCTGCCTCA GCCTCCCTAG
2401 TAGCTAGGAT TACAGGTGCC CGCCACCACA CCCAGCTAGT TTTTGTATTT
2451 TTAATAGAGA CAGGGTTTCA CCGTGTGGA TAGGCTGTTC TTGAACACCT
2501 GACCTCAGGT GATCTGCCTG GCTCGGCTGG AGTGCAGTGG CGTGATCTCA
2551 GCTCACTGCA AGCTCCGCCT CCCGGGTTCA TGCCATTCTC CTGCCTCAGC
2601 CTCCTGAGTA GCTGGGACTA CAGGTGCCCCG CCACCACGCC CCGCTAATTT
2651 TTTTGTATTT TTAGTAGAGA CGGGGTTTCA ACATGTTAGC CAGGATGGTC
2701 TCGATCTCCT GACCTCGTGA TCCGCCCGCC TCAGCCTCCC AAAGTGCTGG
2751 GATTATAGGC GTGAGCCACT GCGCCCGGCC AATTTACTTT TTATTTTATT
2801 TTATTTTATT TTTTGAGACA GGGTCTTGCT CTGTTGCCCA GGCTAGAGTG
2851 CAGTGATACG ATCTTGCTC ACTGCAACCT CTGCTTCTCA GGCTCAACTG
2901 ATCCTCCAC CTCAGCCCC AGGAGCTGGG ACTACAGGTG CATGCCACCA
2951 TGCCCAGCTA ATTTTTTTTG TTTTGTAGTC AGATGAGGTC TTGCCATGTT
3001 GCCCAGACTG CTTATTTTTT TCTAATCAAC TTTTGCCATA AGGACAAGTT
3051 GCTTTCATTG AACTGAGAGT TTTTATTGGT TGCTTACTAA GTAGAAAAGA
3101 ATATTTATTA AGACAGCTTT TTGTCACTTT TAAAAATGAT GTCTTAAGCT
3151 GGGCATAGTG ACTCACATCT ATAATCCAG CACTTGGGGA GGCTGAGGCA
3201 GGTGAACTGC TTGAGCTCAG GAGTTCGAGA CCAGCCTGGG AAACATGGTG
3251 AAACCCCATC TCTACTAAAA ATACAAAAAT TAGTTGGGCA TGGGGTATGT
3301 ACCTGTGGTC CCAGCTACTC AGGGAGGCTG AGGTGGGAGG ATCACTTGAG
3351 CCCTTGAGCC TCAACTTGAG GAAGTTGAGG CTGCAGTGAG CCAAGATCAG
3401 TGCCACTGCA CTCCAGCCTG GGGCGACAGA GCAAGACTCT CTCCAAAAA
3451 AAAAAAAGT CTTAAAAATA GCTGTTTTTG TTTTCCATGT TTGTTTCATA
3501 AATTTTTTTT TTTTTTTTTT TTTTGAGATA GAGTCTCGCT CTATGGCCCA
3551 GGCTGGAGTG CAGTGGCTCA ATCTTGCTC ACTGCAAACT CTACCTCCTG
3601 GGTCCAAGTG ATTCTCCCGC CTCAGCCTTC CGAGTAGCAG GAATTACAAA
3651 CGTGCGCCAC CACACCTGGC TAATTTTTAT ATTTTAAATA GAGATGGGGT
3701 TTGACTATGT TGGCCAGGCT GGTCTTGAAC TCCTGACTTA GTGATCCGCC
3751 TGCCTTGGCC TCCCAAAGTG CTGGGATTAC AGGCGTGAGC CACTGCGTCC
3801 GGCCTAATTT TAAAAGTTTA AAATGGATAA TTTTATTGG CTGTGTGTTT
3851 CATGATTACC AGACTATGTT TCTCTCTCTT GTAGAGGTCC TTTGTCTCC
3901 AATGTTGCTG TCAACATTCA GTCATTTCTC CTTATTTTAC ATGGCAGCAA
3951 ATATGTATGT TTTGTGGAGC TTCTCTTCCA GCATAGTGAA CATTCTGGGT
4001 CAAGAGCAGT TCATGGCAGT GTACCTATCT GCAGGTAATA TGCTTTAATC
4051 TCGGGGCCCTT TGAGAGTATA AGCACTCTAA GCTATCTGCA GAACGGACAA
4101 AGGGAATGAT TACTGCCATA TTCTACACGT AGTGAGTGCT CAGAACATAT
4151 TTGTTTCTCA CAGTGTATGT AGAGAAGGGA GCCACAGATT GGTGGAGATG
4201 TTGCCTTTTC TGTTCATTTT GCTGATTTCT TCTTACATAT GAATTATGTG
4251 GGTATGTTTA ATTTTAAAGT AGGATAAACA GCGGTTAAGT AAGGGTAGT
4301 GTAGAATTTA AGCATGTCAT TTTTGTAAATC TCATCGGGCC TTGATTTTAT
4351 TAGTTTAGGC CCTCCATTTT ATAGATAGTG GTTCCCAGAC TTCCCGGCTG
4401 CCTCAATCTC CTGGGTCTTT GTTAAATAAC CTTAAGCAAG CTCATTTCCC
4451 CCAGTGTGTT CAGTTCACAG AAAGCTTTAA ATCAGAGCTA TACAATATGA
4501 TTGTCAAGAG TGAGTTTGTT CTGCTTCTT TGCAAGAAATG TAGCAGGGAA
4551 CCACTTCCTA GCCATGGTCT TGAAGATGGT ATCGTTTCTT ATTTTCAAGTA

FIGURE 3B

4601 GGAAATTCTC ATGCATGAAT CCAGGTCCCT AGATGCTGCT AACGTGACAG
4651 TTGGTCAAAT TTTACTTACC TCTCTGTTTG TAAAATGTAC TTAATTAAATA
4701 CAATATAAAA ATTAATTTCT AAAATCTCTA CATTTAGAAA CAGTATATCT
4751 GGCAGTTGTG CTGTGATGTA GTGAAAAACA CTAAGCTTGG CGATAGACCC
4801 AGGTTTCAGAT CCTATTTCTA CTACCAGCTG AGTGATGTTG CAAAAATGAC
4851 TAAACCTCAT GATACTTACC TCCTCATGAC AAGGGGTAA AGAAAGGACT
4901 ACATAAAAGC ATCTACCACA AGCCCCAGAG TAGATGCTTA ATTAGTGTTT
4951 ATCGAATACT TATGTGTATC TAGTCCTTCA AAAAAAGAAG CTGAGCATTG
5001 TGTTTGGCTT GTAAGATAAG TGTATAGTTC TTTCCCAAGC ACTAGTTATG
5051 TTGTAGTTAC AGAGGGTCTG TTTCAGATAC ATTAATTCCT GCTCCATAGG
5101 AGGTTTTTAA AAATGAGCCA CGTTGACTCA AATGGCACTG AAGCCAAAGA
5151 GACTTACGGG ATCATCCAGT CTGTTGTCCC ACCCCAGATA TTCTGATTTT
5201 GTGTGTCTGG AGTACAGCCA GAGAATATAC TCTTGGGAAT GAGTCTTCAT
5251 GTTATAGTTG AGGAAAATGG TAACTGAGAA GTGGAGTGAA TGACCGTGTC
5301 GCTCAGCAGA TCATGCAGCA GGTCAGACTT TTCATCCCCT GTAAAGTCGC
5351 TGAAATGATA GGCAGGAGAA GTATTCATGC CCGTACCCTC ACAGTGATCC
5401 AGATTGAAAC CCGACACTGT TTATCTGTGT AGAAATCAGA AATGAAAACC
5451 ATTTTCATGG CTGGATGTGG TGCCGCACGC CTGTAATCCC AGCTACTCAG
5501 GAGGCTGGGG GACAAGAATA ACTTGAACCC GGTAGGCAGA GGTTGCAGTG
5551 AGCCAAAATT GTACCACTGC ACTTCAGCAG CCGGGGCGAA AGAGTGAAAC
5601 TCTGTCTCAA AAAAAAAAAA AAAGAAAAGA AAAAAAAAAAG TAAACCATTT
5651 TTATACCTCA CTAAATTAT TGTAATGTGA CTTGTTTTTC AGGTGTTATT
5701 TCCAATTTTG TCAGTTACGT GGGTAAAGTT GCCACAGGAA GATATGGACC
5751 ATCACTTGGT GCAGTAAGTA TTTCTATTGT AAATTTTTTT TAATTTAATT
5801 TTTAAATTTA CTTTGAAATA AGTTTAGACT TAGAAGAATG TTGTAAAATT
5851 GATAAGTAGG TTCTCATATA CCCTTCACCC TACTGTTAAC TAACATCGAA
5901 ACCAAGAAAT TAACATTGAA ACAATACAGT TGAATAATTT AGAATTTATA
5951 CATTTGTAAA GCTTTGTAAA TGTCCGGCTA TAGCTTTTAA CCATTGGTCA
6001 TATATATATG TTTACCAGAG CAGAGTATAT CTCAGAACAG TAAGTGTGCA
6051 ATCCTCGTAA ACCAGAGAGC CTAATCCAGT ATTGGAAGAT TCTAATTATA
6101 GATTTGAATC TGGTACTTTA TCCTCCTATT TAGTCAATAT TGGAGTGCCT
6151 ACTAGGTGCT ATGCTAGAGC CTGGGGATAA CAGCTGGTGA GCAAGATGAT
6201 CACGATTATT TGTGTTGGTT TTAGAAAGTG GGAACAACA ACAACAAAAA
6251 AGGCTCCTGC CTCAGAGCT CTTATATTCT GGATGCTTAA AAAAATTTTT
6301 CTTAGGCTGG ATGCAGTGGT TTACACCTGT AATCCCAGCA CTTTGGGAGG
6351 CCAAGGTGAG AGGATGAGCC CAAGAATTCG AAACCAGCCC TGGTAACATA
6401 CCAAGATCCT ATCTGTACAA AAAAATTTAA AAAATTAAC TGGGGTGGTG
6451 GCTTATGCCG GTAGTCTCAG CTAATCAGGA GGCTGAGGAA GGAGGATAGC
6501 TTGAGCCTAG GAGGTTGAGG CTGCGGTGAG CTGTGATTGT ACCACTGCAC
6551 CCCAGCCTGG GTGACATAGC AAGACCCTAT CTCAAAAAAA AAATTTTTTT
6601 TTAAGTGTGT TTTGAGGCTG GGTGCAGTGG CTCACACCTG TAATCCCAGC
6651 ACTTTGGGAG GCTGAGGTGG GCAGCTCACT TGAGGTCAGG AGTTCAAGAC
6701 CAGCCTGGTC AACATGGTGA AACCCTGTCC CTCCTGAAAA TACAATAATT
6751 AGCCAGGTGT GGTGTGTCAT GCTTGTAATC CCAGCTACTC GGGAGGCTGA
6801 GGCAGGAGAA TTAATTGAAC CCAGCGGGTA GAGGTTGCAG TGAGCTGAGA
6851 TTGCACCACT GCACTCCAGC CTGGGTGACA GAACAAGACC CTGTCTCACA

FIGURE 3C

6901 GAACAAGACC CTGTCTCAAA GAAAAAAAT TTTTAAAGT GTCTTTTGAG
6951 TTTAATGGCA GATTTCTGGG CACATGGAAA TCTTTATGTA ATATTTCTTT
7001 ACACATTCAG TTTGTACTTA TTAAATACT AATTCATTTA AATGCATTCA
7051 AATAGGGAAT TTCCTATTTA AAGGAACTCT AAAAAGGTCA ATTTTGAAAA
7101 GAATTCTTAT GTAAAATAAC CATTCCCTAA TTTGTATGTT CCCCAAATTT
7151 GTTTACACTT AATTTTCCTA GTGAGGCCTG TGTCTGTCC TGTGACCACA
7201 TGCTTTCTTA AGCCTCCTTT TTTCCCTTCG TGAATGTTT ATTTTCTTTA
7251 TACAATTTCTG CTCTGATATA ATTTATATAT TTCGAATCAT ATTGTCTACC
7301 TCATTCAACA GCTAAGCACC TAATATATGA AGGCAGTGAA GACCACTAGG
7351 ATGAATCAGA GACTCAGAAT TCGAATTTAG CTGGGGAGAA AACATGCACA
7401 CATCTAATAC AACTGAAAG GAATGAGGAT TCTCTAGAGG ACTTTGGGGG
7451 CTCTAAGAGT GAAGAGACCT TTCTAATTAG CTGAAAGGAC CTGCGAGGGC
7501 ATTTTGATGT GCTCTTGAC AGCTGTTGTC CTCATCTTAT AGATAAGAAA
7551 CTGAAGTGCA AACTTAATGA AGTATGGCAG TAAGGTATTT GGAGTTAGAG
7601 TGGGGGTGAA TCCTGGTTCT GCTACTTACG TGTGATTTCT AGGACATATT
7651 ACTGAACTTC TCTGAATTTT AGTTTCCCTT TATAAAATGG GGATAACACC
7701 ATCTATTTCT GAGGTGCAAA GCAAGTACAT TTAGAGTGCT TAGCACAATA
7751 AGAAGCACAT GGTAAGAAAT GTGGACATGG TAGTTCCTGT TCAGTCATCA
7801 AAATCCTACA GCGCCGTGGT AGGATAACAT TATCCCCAAA TATCTTAATG
7851 AATCTGTGAT TAAAATTCAA GGAAATTAAA TCACCAGGTA TAATGGCATT
7901 TTTAATGAGA AATCTGGGAA AAAAACACCA TTAACAAAGT TGTGTTGTTA
7951 CAAAATGTAA AGCGTTAGTC CTCTTGTTT AGTGAGACGT TATAAGATGC
8001 AGGGGACAGC CAGGCACAGT GGCTCACGCC TGTAGGCCCA ACACTTTGGG
8051 AGCCACGGCA GGAAGATCAC TTGAGCCCAG GAGGTTTGAG ACTAGCCTGG
8101 GCAACAAAGT GAGACCCCAT CTCTACAAA AATTTCAAAA TTAAGCCGGG
8151 CATGGTGGCA TGCACCTGTA ATCCTACCTA CTCAGGAGAG GTGGGAGGGT
8201 GGGAGGAATG CCTGAGCCTA GGAGGGTGAG GCTGCTGTGA GCCATGAGCA
8251 TGCCACTGTG CTCCAACCTG GACAACATAG CGAGACCCCA TCTCAAAAAA
8301 AAAAAAGAA AGTTGAATGG GACTGTATAA ATATGTTTGT AAATTACTGT
8351 ATTGGTACTA TCCTGGATAA TTTTAAACT TTTCTGTAGA GACAGGGTCT
8401 CCCTATGTTG CCAAGGCTGG TCTCAAACTC CTGGGCTCAA GTGATCCTCC
8451 TACCTGGGCC TCCCAAAGTG TTGGGATTAC TGGTGTGAGC CACTACACCC
8501 GGCCAATTGT CTTTCTTAT TCAAGTTGAG ATTTTCTGG TTCTTGATAT
8551 GATGAGTGAT TTTTCAGTTG AAGCCTGATC ATTTTAGATA TGATGAGACT
8601 TTGGATCTTA TTGAAATCTG CTGTTTCAGT GGTCTTCCTC TGACACTGTT
8651 CTGATGAGGA GAGGGGGTGC CGTGACTCGT TACTGCTGGG TGTAGGAGTA
8701 GACGTCCAGG TTCCTCACTC AGCCGCCTTT GCCTCCTGAG TGATAGGGGC
8751 TCTTGTCACT GCAGGGCAGG GATGGGAGCT GAGGGCGTGC AGGCTACCTA
8801 GTGTGCCTCT GCTAATGTCG CTGTGGCTAG GAGGAGCAAG GGTGCTTCTT
8851 TCCGCTGACA CCGCTGTGA GCGTATTGG GATGCCTCAT TACAGTGTGG
8901 CAAGGTGGG AGTCTAGGCT CTGCTCAGCC TTTGCTGGGC ACCCGTTTCT
8951 CTAAATATTG TCTAAAAGGT CTCTTTTGCT AGGCTATCTT TTTTGGTCC
9001 TTGACTAGAG AGAACATGTT GAGGGATGAT CGATATGAGG CCAAAAGAAA
9051 GCCCAGGGAA CTCACCACCA CAACATTGAT TGAATCTCAG GCTTCCTAGC
9101 TGGTCCGCTT TCCTCTCTCT TCCTTTCACA GTCCTCTTAC ATTTGTTTCA
9151 TATGTAACAC CCAGGGTCTT TAGCTGTA CT TAGCTTTTGT AAGCAGAGGG

FIGURE 3D

9201 AGCAGATTCA CTTAAATTAT AATACCAAAT AAAGTTAAAA AACATAAGTA
9251 TGATAGATTT GAAGATTATA TAGATACAGA AAAATGTTTG TGAGCCCAGG
9301 CGCAGTGGCT CACAACGTGA ATCCCAGCAC TTTGGGAGGC CGAGGTGGGT
9351 GGATCACTTG AGGCCAGGAG TTCGAAACCA GCCTGGCCAA CATGGTGGAA
9401 CCCCATCTCT ACTAAAAATA CAAAAATTAG CTGGGCATGG TGGTGTGTAC
9451 CTGTTAGTCC CAGCTACTTG GCAGGCTGAG GTGTGAGAAT TAACTTGAAC
9501 CTGGGAGGCG GAGGTTGCAG TGAGATCGTG CCACCGCACT CCAGTTTGGG
9551 CAATAGCGAG ACTCTGTCTC AAAAAATATA TGTTTATGAA ATAAGTAAAA
9601 AAAAATCAGA TGTGCATATT GATTACAGGT ATATAACCAG TACATAAAAA
9651 TATTGATGGA GAACAAAAGA CCTTCACCTC TTCCCATGGA CCCACACCTC
9701 TTAGGTCTGT TGGATCAGGG TTCATGACTC ACTGTACTTA AACTGTGTAT
9751 GAATGTGAGC GTTTTCTGAG AAGAGAAGGG TTCATTTTCA TTAAATTCTT
9801 CTTTCTGACT CGAAAAAGTG AAAAAAGTCT CTCTGCATGG GAGTAAGCCC
9851 AAATATTTGT CAAAAAACAA GTTGTGATTT ATTCAGACAT ATAAATATTT
9901 AAATTTATAT AAAAGCCACA TCGAGAAAAT TCTAGAAGGA TGATGGAAC
9951 GTGTATGTAA TAATTACAAT AAGTTATAAT CACAAAAAAA CCAGCGTTCC
10001 ATGGAATTGT ACAGATAACG ACAATTTTTT TTAACAGATG GAGAATAATC
10051 ATCTATGGAA TAGTAGTTTA GAAGAACTTC ATAGAATTTT TTTTTTTTTT
10101 TTTTTTTTTT TTTTTTGGAG AGGGAGTTTC GTTCTTGTTG CCCAGGCTGG
10151 AGTGCAAAGG TGCGATCTCG GCTCGCTACA ACCTCTGCCT CCCGGGTTCA
10201 AGCGATTCTC CTGCCTCAAC CTCCTGAGTA GCTGGGATTA CAGGCATGCA
10251 CCACCATGCC CAGCTAATTT TGTATTTTTA GCAGAGACTG GGTTTCTTCA
10301 TGTTGGTCAG GCTGGTCTCG AACTCCAGAC CTCAGGTGAT CTGCCCCGCT
10351 CAGCCTCCCA AAGTCCTGGG ATTACAGGTG TAAGCGACTG TGCCTGGCAG
10401 AACTTCATAG AATTTTAATG CTCTTTTATA TCAACTAATC AAATTATATT
10451 TGCTTCATTT TGGGGAAACG TGTAATTTTG ATTTGTTTTG GGGTTTTTTT
10501 GAGATAAAGT GTCACTCTGT CGCCCAGGCT GGAGTACAGT GGCTCAATCT
10551 TGGCTACCA CAACCTCAGC CTTCCGAGTA GCTGGGACTA CAGGCGCCCA
10601 CCACCACGTC TGGCTAATTT TTGTGTTTTT AGTAGAGACG GGGTTTCACT
10651 ATGTTGGCTA GGCTGGTCTT GAACTCCTGA CCTCAGGTGA TCCACCTGCC
10701 TCGGCCCCCTC AGAGTGCTGG GATTACAGGC GTGAGCCACC GTGCCCCGCT
10751 ACAATTATAG TCTCTGCAC AGAAGCCAGC TTGGTCAAAA TTCAGGTCTT
10801 CTTGGGTCCT CCTTTTGAGG AGTGTTTCATG CTGTCCTTCC ATCTTGCACT
10851 TACCCTGACT TCTAAGAATG CAACCCGAGC TTGTTTCCCT GTTGAGGCCA
10901 CTTGGCAGTT ATATGAGGGA CTGGGGACAT CTGAGATCTC TGGGACTCAT
10951 AATAATTTTC TTTAAAGTTT TAGTAATTCC CCAAATGTAA GATAATCTTG
11001 TATTCTGAAG CAACCCGTCA CATAGAAGAC ATTAAGAAAA CATTGATTAA
11051 GAGAGGTAGA TGCTATTTTC CAGAAACAAC CGTTTTTATA TGAAAAGGTA
11101 GGAACCTTTC TTTTAAATGA TAGGGGCTTC TTTCAAAAGT TATTTTGCTC
11151 TTAGGTGTCT TTTTTTTTTT TTTAAACATC TCATTATATA ATAATTAAAA
11201 ACTTATGGGA AAGTTGCAGG GAATAGTACA GAGGACTCCC ATAAAGTCTT
11251 TTTTGTTTGT TTGTTTTGTT TTGTTTTGAG ACAGAGTCTC GCTGTTTTAC
11301 CCAGGCTGGA GTGCAGTGGG ACAATCTCGG CTCCTGCAA CCTCTGCCTC
11351 CCGGGTTCAA GCAATTCTCG GGCCTTAGCA TCCTAAGTAG GTGGGATTAT
11401 AAGCATCCGC CACCACGCCC AGCTAATTTT TTTTTTTTTT TTTTTTTTTT
11451 TATTTTGTAGT AGAGACGGGG TTTTACCACG TTGGTCAGGC TGGTCTCAAA

FIGURE 3E

11501 CTCCTGACCT CAGGTGATCC ACCTGCCTCG GCCTCCAAAA GTGCTGGGAT
11551 TATAGGCGAG AGCCACTGCA CCCAGCCCCA TGAGTCTTT TAAAAAGCA
11601 GGCAACTCAG GTTTACTAGT TAACATGCAA AAAACTGCAC ATATTTAAAG
11651 TTTGGTAAGC TTTGACATGT AGACACCCGT GAAACCATCA CCACACTCAA
11701 GATCATGGAC ATATTCATCC CAAAAGCTTC CTAGTGGTCA CTCCTTCCTG
11751 CCCCTCCTCT ACCCCTGGCG ACAACTTACC TACTTCTACT AAAGATAAAT
11801 TAGTTTGCAA ATGGAACCAT ACAGCATATA CTAGTATTTG TTGTCCTGGC
11851 CTCATTTACT CTGTATAATT ACTTTGAGAC TCATCCATGT TCTGTGTATC
11901 AGTTTATTCC TTTATTATTT TTGAGACAGG GTCTTACTCT GTTGCCCAGG
11951 CAGGAGTGCA GTGGTGCAAT CATAGCTCAC TGTAACCTTG ACCTCCTGGG
12001 CTTAAGGGAT CCTCATGCCT CACAATGTGC TGGGAATTACA GGCCTGAGCC
12051 ACCACACTGG CAATGTTTTG TTTCTTTATG AAGATGAATA AAGATTTTAC
12101 ATGAATTTTT TAAGATGAAA CATGCTTCAT GCATGCAGGT TTCTTTGGGC
12151 GTATTCATGC CCACTCCCTC TGGTTGGAGC TTTGTCAGAG AAGTGTGAGC
12201 AGTTCTTTCC TAGGCCATAG GTGAAAGATG CGCATGACAC GCTTAGCACT
12251 GTCCTTGCGG TTCATGAGGC ACATACATCT TACTGCCCCG TAGTAAAAAT
12301 TCAGTCTTTC CAAGCGATTA CTGTGTGAAG GACATTTAGT TCCTTCACCT
12351 ATTATTGGGG ACATAAGTAA CTGAAAGCTT TGAAGCTTTG TGCTCACCTA
12401 GAAATGTGCA GCATGTAAAC TTTCTAGAAA ATGTGCTGCT CTTTAGACCT
12451 TGTAGCCACT AAGCAGTTGC ATATTGAGTT TCCCATTCTC CCTGCTGTGT
12501 TACTTTGCAG TCTGGTGCCA TCATGACAGT CCTCGCAGCT GTCTGCACTA
12551 AGATCCCAGA AGGGAGGCTT GCCATTATTT TCCTTCCGAT GTTCACGTTT
12601 ACAGCAGGGA ATGTAAGTAT TTTTATGAAG TGCAGTGCTG GGGATAGTGG
12651 TGATGTTTTT ATGTTGAGTG GGTTCCTGCC CTTAAGTTAG AAATGTCAGT
12701 GCTGGAGCAA TCACAGTTGT GCCGCTTGTT TCTTGCTGCC TTTCAGGCCC
12751 TGAAAGCCAT TATCGCCATG GATACAGCAG GAATGATCCT GGGATGGAAA
12801 TTTTTTGATC ATGCGGCACA TCTTGGGGGA GCTCTTTTGT GAATGTAAGT
12851 TTGAGTGTA TTTGATTGCTA AACTGCTTCC TTGGGTCATG CGCTCCTCCT
12901 ACCCCAGCCT CACCCCTACC CCCCATCCCC ATGGCAGAGA CATTGAACTA
12951 TGCAACGGAA GCAGAAGCAG GTGGGCTTGG GAGGGTGAGG AAACCTCAAC
13001 ATGGCTTGCT TTGGGTTTAC CCAGCATACC TGGCTCATTG TAGAGACAGT
13051 CTGTGCCCTT ACCCTACGCT TAACCTTAAG TTGCCCCAAC TGTTGGCCTG
13101 TTATTCCCAG CCCCTCTTA GAAGACTGCA GCCTGGCCCC CAGTCTATGC
13151 TGACATCTTC TTTTCCCCCT TCAGACTTTC CTGCCCTCCT CTCCCCTGCC
13201 TGGCGTCCCA CCCTGCTACC CTGACCTCTG TCTCGCCAGT GCTATTTAGA
13251 CATGCTGAGT TGGCGGAGCC ATTGCTCTGT ATGACTGGAG TAGAGGCCGG
13301 TGACTGCAAA CCAATGTGGA CCACTTACTG AGTACCCGCT GTATGCAGGC
13351 ACCAAGCTAG TTCCCTTATG TTATACTATT ACTACTCCA TTTTACTGAT
13401 GGGAACTGA GGCTCAGACA TCATCTTCCC CAGGCCAAAC AGCTCTTCAA
13451 TAGCAGAGCA GAGCTGTAAA CCCACCTCTA TAAGCCCTTT CCACCCCCAC
13501 CACACCATAT GGAATTGGTT GCTAAACTGC TTCCTTGGGT CACAGCAAAT
13551 GGCATTGTGG TTACAAGACC TTCCACGTGT GCTTCAAACA ATGGGGTTTT
13601 GCCTAGACTA GTGCTTAGTA GTAAGTGTAT CACGGAAACA CGGTCAGGAC
13651 TCTTGCGGTC CATCTGATCG TGGGAGACCC GTCAGCATGA GCTGGATCCC
13701 CTCGGGGCCT GTCTTTTCTT ACATAAATGT TGCTTTTTCG CCTTACTTGG
13751 TTTTTATTTT GTTCCGCGAC AATGGAAAAC TTAATTTTTT TTTTATTAA

FIGURE 3F

13801 AAAGAAAAAT CTATTCTGGC CAGGTGCAGT GGCTCACGCC TGTAATCCCA
13851 GCACTTTGGG AGGCCAAGGC AGGCGGATCA CAAGGTCAGG AGATCGAGAC
13901 CATCCTGGCT AACACAGTGA AACCCCGTCT CTAATAAAAA TACAAAAAAC
13951 TTAGCCGGGC GTGGTGGCGG GCGCCTGTAG TCCAGCTAC TCGGGAGGCT
14001 GAGGCAGGAG AATGGTGTGA ACCCAGAAGG CAGAGCTTGC AGTGAGCCGA
14051 GATCACGCCA CTGCACTCCA GCCTGGGCGA CAAAGTGAGA CTCTGTCTCA
14101 AAAAAAAAAA AAAGAAAAAT CTATTCTAAG TGAAGCAGTT TTTCCAGTA
14151 GGTGGCAGAA CTAAATGCCA TTATGCCATT TATAATTTTA AGTGATTAAA
14201 GAGGAGTAGT ATGTAGTATA TGCAAGGTCT AGCTCTAACA GCAGTGCAGT
14251 ATAAATAGTA GAAACTGACC TGATATTACA GTATGAGAAA CATGAAGGGG
14301 TTCTGTTTTG TGAGCTCTAA ATTTATCTTC CATGTATACT TCAAGGCTCT
14351 TCTCCCGAGT AGATTTTTAT TCATCTGAAC TATAATTAGG TGGCCTTTTT
14401 CCATTCTGAA AATAATTGGA TCAAATGCAT TTTAAAGTCC AGGGTCTGAA
14451 AGGTGGAGGA ATCCTTTCTC TTTACTGTTT CTAATTTAAA CTCCTTTTCA
14501 TTTACTAGAT TTCAGTCATG TCCAGAATTC ATCTTTTCTA AAAGCTTTAA
14551 TCTAGATTTA GAAATCTAAA ATCTTTTATT TATTTTTTTT TCGTTGAAGT
14601 GCCCTGATTT TGTTGGTGGT AAAGACTCCA TTAGTATCCA CTTATACATT
14651 TCCCTGACTT TGCCTCTGAC CAAACCTTAC AGTATTCACA TTGTACTGTT
14701 GCAATAATAA TAGCTAACAT ATTAATACAC TGAATATTTG CTGTGTGCCT
14751 AAGCTAAGGA TTTAATTCTC TTAATAATCCT GTGAGGTATT TTATTTTACA
14801 GAAAAAGAAA CTGCTTAAAG AAAGTAACTT ATCCAGGTCA CACAAGTAAC
14851 AATTGCAGAG CTGGAGTTTC AGATGAGGGC TGGCTTGCGC TGCCGCTACA
14901 GAAAAGAGTG CCCTAGAAAT CGGTCATCTT GCATTTCCCG ATTTTAGTTT
14951 AGCCAAATGA AAAATTCCTT TTGGATTTAT GAGTATAATC AGACAGTATA
15001 CCTGTGAAAT TAAAGTATTT GACTCTTTGC TTGAAATAAG TAGGTTAAAA
15051 AGATTTGGGT GGCCGGGCGC AGTGGCTCAC GCCTGTAATC CCAGCACTTT
15101 GGGAGGCTGA GGCAAGTAGA TCATTTGAGG TCAGGAGTTC GAGACCAGCC
15151 TGACCAATAT GGGGAAACCT CGTCTCTACT AAAAATACAA AAATTAGCCG
15201 GCGGTGGTGG TGCATGCCTG TAATACCAGC TACTTGAGG CTGAGGCAGG
15251 AGAATCACTT GAAGCCAGGA GGCAGAGGTT ACAGTGAGCT GAGATCACGC
15301 CACTGCACTC CAGCCTGGGC AACAGAGCGC GACTCTGTCT AACAAACAAA
15351 AAGATTTGGG AAAACACTTT ATTAATGAAG AGTTCCTGAC AAAGTGATTT
15401 TTTTGGGGAG AATTTTTATA ATTGCATTTG AATATTAGGG TGCTCCTTTT
15451 TCTCTCATT CAAATTCACC AGAGACTTAA GCACAGAGAA TTTTATTAC
15501 ATGCCTGTTA ATTAATGTGT ATAATCAGAT TTAACTATA TTTAGTGAAT
15551 ATTAAGATTC AGGTACAAAT CAAGCCCTTT ATAATTAAAC ATACACATTC
15601 AGAACATTTT TAAAATATTA AAACATTAAA CTGCTCTTCT CACCCACTCC
15651 AAGTCAAATA GCATTTTTTC AGTCAGGTGT CTGGGAGCTC GATGCAAGAT
15701 AACAAAATCT GGTCTCTGCC TCAGGGAACA TGAAATCTGT TTGGGGAAGC
15751 CAGAGCAAAA ATAAAGGTTT TAATAGCAAG CTCTCACTAA CTGCCCCTGG
15801 AAATCCACCC CACATCCTCC AGGAAGCCTT TCTCTACCCC CAGTGCCCTC
15851 AGGAGCTTCT CCAAGGCAGG CCCTTCCCAG AGCGCAGTGT GCTCCCAGC
15901 TCACAGGAGA TGCTCCCTAC ACGCTGCAGG AAAGTCCAGT GCCTGCAGCA
15951 CAGGCTTCAG CAGCAGACTC GGGTTCTAGT CTCAGTCTGC TGATTCTAG
16001 TTGTGGAACC TGAGCAGGCG AAGTTACTAA ACCTCTCTGT GCGTCAGCCT
16051 CCCAGGCTCG TTGCTTCAGG CCGCAGTTAG GCTGTGTGAA CAGGAGAGTG

FIGURE 3G

16101 GGGATGGGAA CTAGGTATCT TAAAGCGGGG CAGAGTTTGG ATGAGCGGGC
16151 CACCCCTTCGT ATAGTTAGGA GGAAGATGAC GGGAGGCATG GAAGCTGGGA
16201 TAGCCATCCT GAGTCAGTGC TAATTCTGAC ACTTCAGAAC ATCGAGTCAG
16251 TCTGACCTGC GAGTGAGCTT TCATTGACCA CTTAGAACT ATTAGCACCT
16301 TGGACAACT ACTTTCTTTC AGACCTGGTT GCTTCATGTC TGCGATGGGA
16351 AAAGTATAC TTAAGTTGCA GATAGTGGTG AATCAAAAGT AGTATATGTG
16401 AAGTACTCAC AACTGCGGA GCATTGAGCC ATCGTCCCAT CCTACTTCTA
16451 CCTTTTACAT ATTGTAATAT GAAAGCTAAA CCATTTCTCG ATGTGAGTCA
16501 GTTTTAATCG GCTACATAGT GAGTGGCATT CGATTTTAAA AATGTCAACT
16551 TGGGATCTGT CACCATGCTA CTTACCATT GTATGTCACA CTGTTTGAAT
16601 GTCGGACCTG GTTTGTTTTT CTCCAGATGG TATGTTACTT ACGGTCATGA
16651 ACTGATTTGG AAGAACAGGG AGCCGCTAGT GAAAATCTGG CATGAAATAA
16701 GGAATAATGG CCCCCAAAAA GGAGGTGGCT CTAAGTAAAA CTGGGATTGG
16751 ACAGTAGTGG TGCATCTGGT CCTTGCCGCC TGAGAGCCCC AGGAGACATC
16801 GGCTAGAGTG ACCATGGCTA TGCTCCCGTC TGAAGATGC CAGCATCTGG
16851 CCTCCCCTG TTTTCAGCTG TGTCCCCAG TCCGTGTCTT TTTAGAATGT
16901 GAATGATGAT AAAGTTGTGA AATAAAGGTT TCTATCTAGT TTGTAAGCAG
16951 ATGTGTGTGT TCTCTCTTTA AGGGGCCGAC ACGGCTCTGG CATTTTGCTT
17001 TGGTTGTTGC ATTGACAGGA CCTGGGGAGA GTGCACCCTG AAAGGCCTGA
17051 TCAGAACATG AAGGCGCTGG TTGCTGTCT TTGGACCCTC CAGTGCCTCT
17101 GCTTAGCCTT CACTCTTCTT TGCTCCCCC TCCCCTGGGT TGGCTGCACA
17151 TAAAAGTCAA GAGTATCCCC TCTCCAGCAC AATCTGAAAT AACAGCTGCA
17201 GTATTTTCTC AATTTTCAGG AAAGGTAGTG TTTTCTGGCA GTGAGTGGCA
17251 TATACAAAAA GCTATTTTCA GGTTTTGCTT TCTAGGTTCA ATTTGTAGAT
17301 AAATTAAGAG GTAGAAAGAA GTGATTTGGG TAAATTCAGA CTTGAAATCT
17351 GAGCCGAATT TTATCTTCTG TTTGAAAGTG TTCTAATTGA AGCGTCTCAC
17401 TGAAAATAGC AGATAGTGGC TGTCGTGTC ACAGCCCTCA CTGTTGTGGA
17451 ATTCATGTTA CCCTCGTGAC TGAGAATGAC ATCTAGGAAA TGCAGTTTGA
17501 GAGTATGTTT TTCTTGAAGT CATTTACAGG AGAATTTTGA GTCTTTTGAT
17551 GGCTTCAAAA TGTTATACCA AGTCTTGCGC CTTTGTCTCT GGAGGATCGA
17601 AGGCCCTGAT TTCAGCCTCC TGTGGCCGAT CGGACTCAGG TTGTGTGCCG
17651 TGGGGGATGG GAATGGCGGC TTTGGAAAAG GAGTGGGAGT GGTGCCACC
17701 TCACCAGGCA AGTGAGAACT GCATGGCAGC ACGCGCCCAG CACATAGAAA
17751 TTGTCCAGTA TTTGGCAGTC CTTCATATCC TTCTTCCATC AGGCTGGACT
17801 TGTTTCTACT ATGATTTACA GTTATTCTTC CCAGGCACAG GATTCTGTTT
17851 TAAACTCGTA TCACTTCTAG GGGAGAGAGT TATCTTAGCC ATCATTTTGC
17901 CAGCGAGGAA ACGGCACACG TGGTGTAGGG GCACTGCCCA AGGTCACAAT
17951 GCTTTGCTCT GACATCTGCT AACAACTGCA ACACAGATGA GGCAAGATGC
18001 GTTTTCCAGA GATGGGATAG GAGGCTGAGT TCATAGGGAC ATTCCCTCTA
18051 GAGCCCAACA TTAATTCACA TCGTGCTTTG GGCAGACCAG GCAAAGAGGC
18101 AATGAAGACA TCTCTGTGTC CCTGCTTTGT GACTGGGAAA AAGTTAGAAG
18151 TCCCTGTAGC ATCTCCTGGT CCCTAAAACC CCTCAATGCT GGAGCCTCTG
18201 TGCATGGCCT GGGGAGGCCA GAACCTGGCT GTGGCCGGAG AAGCCTTGCT
18251 GTCCACAGCT CCCTCCTGAT TGCCACGAG GGTGCTTAC TTTCTCCTCT
18301 TGGCTTCTCT GGGGACCCGC GATCACTGCC TTCAAGGCCA TGCACTCCCT
18351 GGCCCGTGGG CCTCTTGGGC TGTGCCGCCT CCACTGGCAT CTGAAGTGTG

FIGURE 3H

```

18401 GGGTACCTAG GAACATGCCG TGGCTGCCGT CTCCTCATT CCATACACTT
18451 CTTGAGTGGG TGCACCTGCT GAAGCCTCAG TTATCTGTGA GGATTCTGAG
18501 CTCCAGACCC ACAGAATCTC TCTGTACTCT TAGTAAATGT GTCTACTGCA
18551 ACACACGCAT GGTTCAGGC TCTGGGACCA CCCCCCGCC CTGCACAGGC
18601 CCCTCAAATA GCACTCGGCT TAAGGAGTGA CACGAGCAAT CGGTGAAGTC
18651 TGAAACCCGG AGCCATTGCA GATCTCCCTC TCTCGCCTCT TATTTCTAGA
18701 ATTCAGCCCC TCAGCCTTCC CAGTGCCTGT GACTCCGTGG TGGTCCTCAC
18751 TTCTTAGTCC CTGGACTGTT GAGCCTGTTT TTCCAGCTGG TCTCCAAAGC
18801 AACCTGTGC TTCTCCATAT GCCTGCCAGA GTGCTAAAAA CACGTCTGTC
18851 ATTCCTTTGT TGTACCTGT GAAAACTTT TATTTATTTG AGACAGGGTC
18901 TCTCTCTCTC TCTCTCGTCC AGGCTGGAGT TCAGTGGTGC AATCTAGATG
18951 GTCACACAC TCAGGGAGTT GGGGATGGCT CAGAGCTGTT AACAGAGAGG
19001 GGAAGTCCCA GGAGGACCTG CGTGAGGGGT GGGGGTGGGA TGACAAGGAA
19051 CCAGCTCTGG GAGTTGAAAG ACCTGGATTG AAGTCTCAAC CCAAGCCCTG
19101 GCCAGCTCTG GGACCCCGGA CAAGTCGGCC TCACTCTCTG CCCCTCAGTG
19151 GGCTCCTGTG TAGATGGGGA TAATGATGGC TTTATATCCT GAGAATGTGG
19201 GGAGGGGATT AAGTGGCAA AATACCTGAG AGTGCCTGCT CAGTGCCTGG
19251 CTCAGCAAAT GCCCTTGTTT CCTCCTTCCC TCTCCCAGA ACCCTCCTC
19301 CCCTTCTTCT TCTTTTTTTT TTTTTTTTTT TGACCCAGAG TCTTGCTATG
19351 TTGCCCAGGC TGGAGTGCAG TGGCACAATC TCGGCTCACT GCAACCTCCA
19401 CCTCCTGGCT TCAGGCAATT CTTGTGCCTC AGCCTCTCGA GTAGCTGGGA
19451 TTACAGGCAG GCACCATCAC GCCCGGCTAA TTTTTTTTTT TTTTTTTTGT
19501 AGTAGAAATG GGATTTTACC ATATTGGCAG GATGTTCTCG ATCTCCTGAC
19551 CTCAGGTGAT CCACTCGCCT TGGCCTCCA AAGTGCTGGG ATTATAGGTG
19601 TCAGCCACTG CGCCAGCCC CCATTGTTTA TCTCCTCTC CATTTCTTGT
19651 GGGGACTTTT AAAGGAAAAA TCAGTTTGGT GGGCTGGGGG AGGGCATAGC
19701 TGAGACCACC TTGAGGGCAC CAAGCTCACT GACCAC (SEQ ID NO:3)

```

FEATURES:

Start: 2002
Exon: 2002-2098
Intron: 2099-5692
Exon: 5693-5763
Intron: 5764-12510
Exon: 12511-12612
Intron: 12613-12746
Exon: 12747-12844
Intron: 12845-16626
Exon: 16627-16735
Stop: 16736

SNPs:

DNA	Protein					
Position	Major	Minor	Domain	Position	Major	Minor

237	T	C	Beyond ORF(5')
783	G	T	Beyond ORF(5')
1187	C	T	Beyond ORF(5')
1227	-	A T	Beyond ORF(5')
1450	T	C G	Beyond ORF(5')
3925	C	T	Intron
5539	G	C	Intron
7220	T	C	Intron
7396	G	A	Intron
9048	A	C	Intron
9952	T	C	Intron
10197	G	A T	Intron
10245	C	G	Intron
10427	C	T	Intron
10583	T	C	Intron
10651	A	G	Intron
11125	G	A	Intron
12025	A	C	Intron
12391	T	G	Intron
13001	A	G	Intron
13147	A	G	Intron
13587	A	G	Intron
13681	T	G	Intron
14336	A	G	Intron
14729	A	G	Intron
15124	C	T	Intron
15907	A	G	Intron
16341	-	G T	Intron
16786	G	C	Beyond ORF(3')
17159	G	A	Beyond ORF(3')
17976	-	T C	Beyond ORF(3')
18001	G	A	Beyond ORF(3')
18021	G	T	Beyond ORF(3')
18022	A	G	Beyond ORF(3')
18042	T	G	Beyond ORF(3')
18375	C	T	Beyond ORF(3')
19244	T	C	Beyond ORF(3')

Context:
DNA
Position

237 CGAGGTTTCTTCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGGTGATCCGTCCG
CCTCAGCCTCCCAAAGTACTGCTGGGATTACAGACGTGAGCCACCGCACCCGGCCTTTAT
CTTTCATTTTTTTTCATGTATTTTCCTTTATTTTAATCACTTTATCCAGAAACATATCCT

FIGURE 3J

CGTCTTGACAGTGCTGTGGTGCCTGTGGTTTCCAGAAGCTGGGTGTGCTGTGTGTC
[T,C]

GTGGTTTGAGGAAGTTGCCCATGGAAGTACAGAGGAAGCAGAGTAGTCGTTGCCATTTT
TCAGCCTAGTAGGCAGGATCAGGGACCCCATCTTGCTCTCTTTGCCTTGAACCACAATTA
GAATAAAACACCAAAGCCCTGACTGATCATGATCATAGCAATCCGATCTTTATGATCATG
GCCAGACCATTCTCAGGTCGTCTTTACCCTAAGATATCAATCACTGGGTATGACAACCTA
GACCTAAGGGTGCACTCTGGGTAGTAAAGATGATTAAGTCTCCCAAAGGAATCTAAGGAA

783

AAGGGTGCACTCTGGGTAGTAAAGATGATTAAGTCTCCCAAAGGAATCTAAGGAATCCAG
AGCAACACGAATCACTGCTCTCTTCTATAGGGTAAACCTCCCAAGACTCCAGTCCCTGT
GAGGAGGCTCTGCCCCGCTGCCCTTCCCAGGGTTCCAGGCTCCACATTGGGAGGTGTACA
CAGTGCTCTTCGCTCTTCATTGCCTTGTGTATGATCCCTTTTCCCATCTTTCATAAATG
CTGTCCCTCTCACCATCTTTAAAAGAGTCTGGGTAATTATTTACCAAAGGTGGTATAAT
[G,T]

CTGTCACAGTCCCTGCTAGTGAGACATCTGATACAACTGATGGAATCAGTTCAACAAAAT
GCAGTAAATTTTATTTAATGTACTACGGAGAAAGAAAAATGCTACCAGTTATAAGATG
CATCCTGATTTTCAGATATTTAAATGGAAAAATGTCTTAAGATCTGTGAAAAATGTAGCT
TCCTTTCCACCTCTCAAGTGGGAGAGCAAAAACTGGACAGACTAGAAATGCCAGGGGCT
AGCTGAGAACCTTACAGAATGAGCAACTGCGGAAGCCACAGGTAAACCCGAGATGTAGAT

1187

CTACCAGTTATAAGATGCATCCTGATTTTCAGATATTTAAATGGAAAAATGTCTTAAGAT
CTGTGAAAAATGTAGCTTCCTTTCCACCTCTCAAGTGGGAGAGCAAAAACTGGACAGAC
TAGAAATGCCAGGGGCTAGCTGAGAACCTTACAGAATGAGCAACTGCGGAAGCCACAGGT
AACACCGAGATGTAGATCAGCTGCCAGGGACAAGACAAAGAATGTTTTCTAAAGTAAATC
CTCTTACCAGTATGTTATTGAAATCAGTCCTTATTGGCATCGAAGAAGGTGAAAGTGCTA
[C,T]

TTGCCTGTTGCCTACAGAGACTGGAGGAATGACAAATGTTTAAATTATTTTAATTCAACA
AGTAGAGGAATACCTGCTATGTGAAGGAGTTGTGGCAATTCATAAAATTAATATATTTT
TGAAGTTTGTAGTTTTCAATAATAATTTCTTATCTAAAATGTAACAAGTTAATTATATTA
TCGAATAAACCTCAATTTTCGTAGTACTAACAACATCAACACTTACAGAAAAAGGAAAGTC
ACTCAACTCCCATGTAAACAGACTTTAGAAGCAGTTGCAGAGGTTTTCTAAATTATCC

1227

TGGAAAAATGTCTTAAGATCTGTGAAAAATGTAGCTTCCTTTCCACCTCTCAAGTGGG
AGAGCAAAAACTGGACAGACTAGAAATGCCAGGGGCTAGCTGAGAACCTTACAGAATGAG
CAACTGCGGAAGCCACAGGTAAACCCGAGATGTAGATCAGCTGCCAGGGACAAGACAAAG
AATGTTTTCTAAAGTAAATCCTCTTACCAGTATGTTATTGAAATCAGTCCTTATTGGCAT
CGAAGAAGGTGAAAGTGCTACTTGCCTGTTGCCTACAGAGACTGGAGGAATGACAAATGT
[-,A,T]

TAAATTATTTTAATTCAACAAGTAGAGGAATACCTGCTATGTGAAGGAGTTGTGGCAATT
CATAAAATTAATATATTTTTTGAAGTTTGTAGTTTTCAATAATAATTTCTTATCTAAAAT
GTAACAAGTTAATTATATTATCGAATAAACCTCAATTTTCGTAGTACTAACAACATCAACA
CTTACAGAAAAAGGAAAGTCACTCAACTCCCATGTAAACAGACTTTAGAAGCAGTTGC
AGAGGTTTTCTAAATTATCCCTGAATTCCTATCATGACTATTTTTCTCAGACATGTTG

1450

TCAGTCCTTATTGGCATCGAAGAAGGTGAAAGTGCTACTTGCCTGTTGCCTACAGAGACT
GGAGGAATGACAAATGTTTAAATTATTTTAATTCAACAAGTAGAGGAATACCTGCTATGT

FIGURE 3K

GAAGGAGTTGTGGCAATTCATAAAATTAATATATTTTTTTGAAGTTTGTAGTTTTCAATAA
TAATTTCTTATCTAAAATGTAACAAGTTAATTATATTATCGAATAAACCTCAATTTTCGTA
GTAATAACAACATCAACACTTACAGAAAAAGGAAAGTCACTCAACTCCCACATGTAAACA
[T, C, G]

ACTTTAGAAGCAGTTGCAGAGGTTTTCTAAATTATCCCTGAATTCCTATCACATGACTAT
TTTTCTCAGACATGTTGACCTTACCTACACAGATGACTCACATATGTTTCCATAAGCTG
GCAGTAAGTTTAAGAAGCATACCATGCCCTGAGGAAAAAGAAGTAATGTTAGCTCTTCTA
CTCTTGGCCAAAGAACCTAATTCTGTATATTACTTCTGTCTTTGGTTTGGCTATTATAGA
CAATAAATTATTGATCTGATTATAATTGAGAAAAGTAAGCTCTTCTAAAGAAGTAAAATA

3925 GCCTTCCGAGTAGCAGGAATTACAAACGTGCGCCACCACACCTGGCTAATTTTTATATTT
TTAATAGAGATGGGGTTTGACTATGTTGGCCAGGCTGGTCTTGAACTCCTGACTTAGTGA
TCCGCCTGCCTTGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACTGCGTCCGGCC
TAATTTTAAAAGTTTAAATGGATAATTTTTATTGGCTGTGTGTTTCATGATTACCAGAC
TATGTTTCTCTCTCTTGTAGAGGTCTTTGTTCTCCAATGTTGCTGTCAACATTCAGTCA
[C, T]

TTCTCCTTATTTACATGGCAGCAAATATGTATGTTTTGTGGAGCTTCTCTTCCAGCATA
GTGAACATTCTGGGTCAAGAGCAGTTTCATGGCAGTGTACCTATCTGCAGGTAATATGCTT
TAATCTCGGGGCTTTGAGAGTATAAGCACTCTAAGCTATCTGCAGAACGGACAAAGGGA
ATGATTACTGCCATATTCTACACGTAGTGAGTGCTCAGAACATATTTGTTTCTCACAGTG
TATGTAGAGAAGGGAGCCACAGATTGGTGGAGATGTTGCCTTTTCTGTTCAATTTTGCTGA

5539 ATGAGTCTTCATGTTATAGTTGAGGAAAATGGTAACTGAGAAGTGGAGTGAATGACCGTG
TCGCTCAGCAGATCATGCAGCAGGTGAGACTTTTCATCCCCTGTAAAGTCGCTGAAATGA
TAGGCAGGAGAAGTATTCATGCCCGTACCCTCACAGTGATCCAGATTGAAACCCGACACT
GTTTATCTGTGTAGAAATCAGAAATGAAAACCATTTTCATGGCTGGATGTGGTGCCGCAC
GCCTGTAATCCCAGCTACTCAGGAGGCTGGGGGACAAGAATAACTTGAACCCGGTAGGCA
[G, C]

AGGTTGCAGTGAGCCAAAATTGTACCACTGCACCTCAGCAGCCGGGGCGAAAGAGTGAAA
CTCTGTCTCAAAAAAAAAAAAAAAAAAGAAAAAGAAAAAAAAAAGTAAACCATTTTTATACCTC
ACTTAAATTATTGTAATGTGACTTGTTTTTCAGGTGTTATTTCCAATTTTGTGAGTTACG
TGGGTAAAGTTGCCACAGGAAGATATGGACCATCACTTGGTGCAGTAAGTATTTCTATTG
TAAATTTTTTTTAAATTTAATTTTTAAATTTACTTTGAAATAAGTTTAGACTTAGAAGAAT

7220 AGAAAAAAAAATTTTTTAAAGTGTCTTTTGAGTTTAAATGGCAGATTTCTGGGCACATGGAA
ATCTTTATGTAATATTTCCCTTACACATTCAGTTTGTACTTATTTAAATACTAATTCATTT
AAATGCATTCAAATAGGGAATTTCCCTATTTAAAGGAAGTCTAAAAAGGTCAATTTTGAAA
AGAATTCTTATGTAAATAACCATTCCTAATTTGTATGTTCCCAAATTTGTTTACACT
TAATTTTCCCTAGTGAGGCCTGTGTTCTGTCTGTGACCACATGCTTCTTAAGCCTCCTT
[T, C]

TTTCCCTTCGTGGAATGTTTATTTTCTTTATACAATTTGCTCTGATATAATTTATATAT
TTCGAATCATATTGTCTACCTCATTCAACAGCTAAGCACCTAATATATGAAGGCAGTGAA
GACCACTAGGATGAATCAGAGACTCAGAATTCGAATTTAGCTGGGGAGAAAAACATGCACA
CATCTAATACACACTGAAAGGAATGAGGATTCTAGAGGACTTTGGGGGCTCTAAGAGT
GAAGAGACCTTTCTAATTAGCTGAAAGGACCTGCGAGGGCATTTTGATGTGCTCTTGGAC

FIGURE 3L

7396 GAAAAGAATTCTTATGTAAAATAACCATTCCCTAATTTGTATGTTCCCAAATTTGTTTAA
CACTTAATTTTCTAGTGAGGCCTGTGTTCTGTCTGTGACCACATGCTTTCTTAAGCCT
CCTTTTTTCCCTTCGTGGAATGTTTATTTCTTTATACAATTCGCTCTGATATAATTTA
TATATTTTGAATCATATTGTCTACCTCATTCAACAGCTAAGCACCTAATATATGAAGGCA
GTGAAGACCACTAGGATGAATCAGAGACTCAGAATTCGAATTTAGCTGGGGAGAAAACAT
[G,A]
CACACATCTAATACACACTGAAAGGAATGAGGATTCTCTAGAGGACTTTGGGGGCTCTAA
GAGTGAAGAGACCTTTCTAATTAGCTGAAAGGACCTGCGAGGGCATTTTGATGTGCTCTT
GGACAGCTGTTGTCCTCATCTTATAGATAAGAACTGAAGTGCAAACTTAATGAAGTATG
GCAGTAAGGTATTTGGAGTTAGAGTGGGGGTGAATCCTGGTTCTGCTACTTACGTGTGAT
TTCTAGGACATATTACTGAACCTCTCTGAATTCAGTTTCCCTTTATAAAATGGGGATAA

9048 GGCTCTTGCTCACTGCAGGGCAGGGATGGGAGCTGAGGGCGTGCAGGCTACCTAGTGTGCC
TCTGCTAATGTCGCTGTGGCTAGGAGGAGCAAGGGTGCTTCTTTCCGCTGACACCGCCTG
TTAGGCGTATTGGGATGCCTCATTACAGTGTGGCAAGGGTGGGAGTCTAGGCTCTGCTCA
GCCTTTGCTGGGCACCCGTTTCTCTAAATATTGTCTAAAAGGTCTCTTTGCTAGGCTAT
CTTTTTTGGTCCTTGACTAGAGAGAACATGTTGAGGGATGATCGATATGAGGCCAAAAG
[A,C]
AAGCCAGGGAACTCACCACCACAACATTGATTGAATCTCAGGCTTCCTAGCTGGTCCGC
TTTCTCTCTCTTCTTTTACAGTCTCTTACATTTGTTTCATATGTAAACCCAGGGTC
TTTAGCTGTACTTAGCTTTTGTAAAGCAGAGGGAGCAGATTCACTTAAATTATAATACCAA
ATAAAGTTAAAAAACATAAGTATGATAGATTTGAAGATTATATAGATACAGAAAAATGTT
TGTGAGCCCAGGCGCAGTGGCTCACAACTGTAATCCCAGCACTTTGGGAGGCCGAGGTGG

9952 ATTGATGGAGAACAAAAGACCTTCACCTCTTCCCATGGACCCACACCTCTTAGGTCTGTT
GGATCAGGGTTCATGACTCACTGTACTTAACTGTGTATGAATGTGAGCGTTTTCTGAGA
AGAGAAGGGTTCATTTTCATTAAATTCCTCTTCTGACTCGAAAAAGTGAAAAAGTCTC
TCTGCATGGGAGTAAGCCCAAATATTTGTCAAAAAACAAGTTGTGATTTATTAGACATA
TAAATATTTAAATTTATATAAAAGCCACATCGAGAAAATTCTAGAAGGATGATGGAAGT
[T,C]
GTATGTAATAATTACAATAAGTTATAATCACAAAAAAACCAGCGTTCATGGAATTGTAC
AGATAACGACAATTTTTTTTAAACAGATGGAGAATAATCATCTATGGAATAGTAGTTTGA
AGAACTTCATAGAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGAGAGGGAGTTTCGT
TCTTGTGCCCAGGCTGGAGTGCAAAGGTGCGATCTCGGCTCGCTACAACCTCTGCCTCC
CGGGTTCAAGCGATTCTCCTGCCTCAACCTCCTGAGTAGCTGGGATTACAGGCATGCACC

10197 ATTTAAATTTATATAAAAGCCACATCGAGAAAATTCTAGAAGGATGATGGAAGTGTGTAT
GTAATAATTACAATAAGTTATAATCACAAAAAAACCAGCGTTCATGGAATTGTACAGAT
AACGACAATTTTTTTTAAACAGATGGAGAATAATCATCTATGGAATAGTAGTTTGAAGAA
CTTCATAGAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGAGAGGGAGTTTCGTCTT
GTTGCCAGGCTGGAGTGCAAAGGTGCGATCTCGGCTCGCTACAACCTCTGCCTCCCGGG
[G,A,T]
TCAAGCGATTCTCCTGCCTCAACCTCCTGAGTAGCTGGGATTACAGGCATGCACCACCAT
GCCAGCTAATTTTGTATTTTAGCAGAGACTGGGTTTCTTCATGTTGGTCAGGCTGGTC
TCGAACTCCAGACCTCAGGTGATCTGCCCGCTCAGCTCCCAAAGTCTGGGATTACAG
GTGTAAGCGACTGTGCCTGGCAGAACTTCATAGAATTTAATGCTCTTTTATATCAACTA

FIGURE 3M

ATCAAATTATATTTGCTTCATTTTGGGGAAACGTGTAATTTTGATTGTGTTTGGGGTTTT

10245 GGAAGTGTGTATGTAATAATTACAATAAGTTATAATCACAAAAAACAGCGTTCATGG
AATTGTACAGATAACGACAATTTTTTTTAAACAGATGGAGAATAATCATCTATGGAATAGT
AGTTTAGAAGAACTTCATAGAATTTTTTTTTTTTTTTTTTTTTTTTTTTTGGAGAGGG
AGTTTCGTTCTTGTTGCCAGGCTGGAGTGCAAAGGTGCGATCTCGGCTCGCTACAACCT
CTGCCTCCCGGTTCAAGCGATTCTCCTGCCTCAACCTCCTGAGTAGCTGGGATTACAGG
[C,G]
ATGCACCACCATGCCCAGCTAATTTTGTATTTTAGCAGAGACTGGGTTTCTTCATGTTG
GTCAGGCTGGTCTCGAACTCCAGACCTCAGGTGATCTGCCCCGCTCAGCCTCCCAAAGTC
CTGGGATTACAGGTGTAAGCGACTGTGCCTGGCAGAACTTCATAGAATTTTAATGCTCTT
TTATATCAACTAATCAAATTATATTTGCTTCATTTTGGGGAAACGTGTAATTTTGATTGT
TTTTGGGGTTTTTTTGTAGATAAAGTGTCACTCTGTGCCCCAGGCTGGAGTACAGTGGCTC

10427 TTTGTTCTTGTTGCCAGGCTGGAGTGCAAAGGTGCGATCTCGGCTCGCTACAACCTCT
GCCTCCCGGTTCAAGCGATTCTCCTGCCTCAACCTCCTGAGTAGCTGGGATTACAGGCA
TGCACCACCATGCCCAGCTAATTTTGTATTTTAGCAGAGACTGGGTTTCTTCATGTTGG
TCAGGCTGGTCTCGAACTCCAGACCTCAGGTGATCTGCCCCGCTCAGCCTCCCAAAGTCC
TGGGATTACAGGTGTAAGCGACTGTGCCTGGCAGAACTTCATAGAATTTTAATGCTCTTT
[C,T]
ATATCAACTAATCAAATTATATTTGCTTCATTTTGGGGAAACGTGTAATTTTGATTGT
TTGGGGTTTTTTTGTAGATAAAGTGTCACTCTGTGCCCCAGGCTGGAGTACAGTGGCTCAA
TCTTGGCTCACCACAACCTCAGCCTTCCGAGTAGCTGGGACTACAGGCGCCACCACCAC
GTCTGGCTAATTTTGTGTTTTTAGTAGAGACGGGGTTTCACTATGTTGGCTAGGCTGGT
CTTGAACCTCCTGACCTCAGGTGATCCACCTGCCTCGGCCCTCAGAGTGCTGGGATTACA

10583 AGAGACTGGGTTTCTTCATGTTGGTCAGGCTGGTCTCGAACTCCAGACCTCAGGTGATCT
GCCCCGCTCAGCCTCCCAAAGTCTGGGATTACAGGTGTAAGCGACTGTGCCTGGCAGAA
CTTCATAGAATTTAATGCTCTTTATATCAACTAATCAAATTATATTTGCTTCATTTTG
GGGAAACGTGTAATTTTGATTGTGTTTTGGGGTTTTTTTGTAGATAAAGTGTCACTCTGTG
CCCAGGCTGGAGTACAGTGGCTCAATCTTGGCTCACCACAACCTCAGCCTTCCGAGTAGC
[T,C]
GGGACTACAGGCGCCACCACCACGTCTGGCTAATTTTGTGTTTTAGTAGAGACGGGG
TTTCACTATGTTGGCTAGGCTGGTCTTGAACCTCCTGACCTCAGGTGATCCACCTGCCTCG
GCCCCCTCAGAGTGCTGGGATTACAGGCGTGAGCCACCGTGCCCGGCTACAATTATAGTCT
CTTGACAGAAAGCCAGCTTGGTCAAAATTCAGGTCTTCTTGGGTCTCTTTTGAGGAGT
GTTGATGCTGTCTTCCATCTTGACAGTACCCTGACTTCTAAGAATGCAACCCGAGCTTG

10651 CAGCCTCCCAAAGTCTGGGATTACAGGTGTAAGCGACTGTGCCTGGCAGAACTTCATAG
AATTTAATGCTCTTTTATATCAACTAATCAAATTATATTTGCTTCATTTTGGGGAAACG
TGTAATTTTGATTGTGTTTTGGGGTTTTTTTGTAGATAAAGTGTCACTCTGTGCCCCAGGCT
GGAGTACAGTGGCTCAATCTTGGCTCACCACAACCTCAGCCTTCCGAGTAGCTGGGACTA
CAGGCGCCACCACCACGTCTGGCTAATTTTGTGTTTTAGTAGAGACGGGGTTTCACT
[A,G]
TGTTGGCTAGGCTGGTCTTGAACCTCCTGACCTCAGGTGATCCACCTGCCTCGGCCCTCA
GAGTGCTGGGATTACAGGCGTGAGCCACCGTGCCCGGCTACAATTATAGTCTCTTGACA

FIGURE 3N

11125
TTCATGCTGTCCTTCCATCTTGCAGTTACCCTGACTTCTAAGAATGCAACCCGAGCTTGT
TTCCCTGTTGAGGCCACTTGGCAGTTATATGAGGGACTGGGGACATCTGAGATCTCTGGG
ACTCATAATAATTTTCTTTAAAGTTTTAGTAATCCCCAAATGTAAGATAATCTTGTATT
CTGAAGCAACCCGTACATAGAAGACATTAAGAAAAACATTGATTAAGAGAGGTAGATGCT
ATTTTCCAGAAACAACCGTTTTTATATGAAAAGGTAGGAACCTTCTTTTTAATGATAGG
[G,A]
GCTTCTTTCAAAGTTATTTTGCTCTTAGGTGTCTTTTTTTTTTTTTTAAACATCTCATT
CATAAATAATTA AAAA ACTTATGGGAAAGTTGCAGGGAATAGTACAGAGGACTCCCATAAA
GTCTTTTTTGTTTGTTTGTTTGTTTGTTTGTTTGAGACAGAGTCTCGCTGTTTTACCCAGG
CTGGAGTGCAGTGGGACAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGCAAT
TCTCGGGCCTTAGCATCCTAAGTAGGTGGGATTATAAGCATCCGCCACCACGCCCAGCTA

12391
AAGATTTCACATGAATTTTTTAAGATGAAACATGCTTCATGCATGCAGGTTTCTTTGGGC
GTATTCATGCCCCACTCCCTCTGGTTGGAGCTTTGTGAGAGAAGTGTGAGCAGTTCTTTCC
TAGGCCATAGGTGAAAGATGCGCATGACACGCTTAGCACTGTCCTTGCGGTTTCATGAGGC
ACATACATCTTACTGCCCCGTAGTAAAAATTGAGTCTTTCCAAGCGATTACTGTGTGAAG
GACATTTAGTTCCTTCACCTATTATTGGGGACATAAGTAACTGAAAGCTTTGAAGCTTTG
[T, G]
GCTCACCTAGAAATGTGCAGCATGTAACTTTCTAGAAAATGTGCTGCTCTTTAGACCTT
GTAGCCACTAAGCAGTTGCATATTGAGTTTCCATTCTCCCTGCTGTGTTACTTTGCAGT
CTGGTGCCATCATGACAGTCCTCGCAGCTGTCTGCACTAAGATCCCAGAAGGGAGGCTTG
CCATTATTTTCTTCCGATGTTTACGTTTACAGCAGGGAATGTAAGTATTTTTATGAAGT
GCAGTGCTGGGGATAGTGGTGATGTTTTATGTTGAGTGGGTTCTTGCCCTTAAGTTAGA

FIGURE 30

TGGCTTGCTTTGGGTTTACCCAGCATACCTGGCTCATTGTAGAGACAGTCTGTGCCTTTA
CCCTACGCTTAACCTTAAGTTGCCCAACTGTTGGCCTGTTATTCCCAGCCCCCTCTTAG
AAGACTGCAGCCTGGCCCCAGTCTATGCTGACATCTTCTTTTCCCCTTCAGACTTTCC
TGCCCTCTCTCCCCTGCCTGGCGTCCCACCCTGCTACCCTGACCTCTGTCTCGCCAGTG
CTATTTAGACATGCTGAGTTGGCGGAGCCATTGCTCTGTATGACTGGAGTAGAGGCCGGT

13147 AAGTTTGAGTGTAATTGATTGCTAAACTGCTTCCTTGGGTGATGCGCTCCTCCTACCCCA
GCCTACCCCCTACCCCCATCCCCATGGCAGAGACATTGAACTATGCAACGGAAGCAGAA
GCAGGTGGGCTTGGGAGGGTGAGGAAACCTCAACATGGCTTGCTTTGGGTTTACCCAGCA
TACCTGGCTCATTGTAGAGACAGTCTGTGCCTTTACCCTACGCTTAACCTTAAGTTGCC
CAACTGTTGGCCTGTTATTCCCAGCCCCCTCTTAGAAGACTGCAGCCTGGCCCCCAGTCT
[A,G]

TGCTGACATCTTCTTTTTCCCCTTCAGACTTTCTGCCCTCCTCTCCCCTGCCTGGCGTC
CCACCCTGCTACCCTGACCTCTGTCTCGCCAGTGCTATTTAGACATGCTGAGTTGGCGGA
GCCATTGCTCTGTATGACTGGAGTAGAGGCCGGTGACTGCAAACCAATGTGGACCACTTA
CTGAGTACCCGCTGTATGCAGGCACCAAGCTAGTTCCCTTATGTTATACTATTACTACTC
CCATTTTACTGATGGGAAACTGAGGCTCAGACATCATCTTCCCAGGCCAAACAGCTCTT

13587 GGAGTAGAGGCCGGTGACTGCAAACCAATGTGGACCACTTACTGAGTACCCGCTGTATGC
AGGCACCAAGCTAGTTCCCTTATGTTATACTATTACTACTCCCATTTTACTGATGGGAAA
CTGAGGCTCAGACATCATCTTCCCAGGCCAAACAGCTCTTCAATAGCAGAGCAGAGCTG
TAAACCCACCTCTATAAGCCCTTTCACCCCCACCACCATATGGAATTGGTTGCTAAA
CTGCTTCCTTGGGTACAGCAAATGGCATTGTGGTTACAAGACCTTCCACGTGTGCTTCA
[A,G]

ACAATGGGGTTTTGCCTAGACTAGTGCTTAGTAGTAACTGTATCACGGAAACACGGTCAG
GACTCTTGGCGTCCATCTGATCGTGGGAGACCCGTCAGCATGAGCTGGATCCCCTCGGGG
CCTGTCTTTTCTTACATAAATGTTGCCTTTGCCCCTTACTTGGTTTTTATTTTGTTCGC
GACAATGGAAAACCTAATTTTTTTTTTATTAAGAAAGAAAAATCTATTCTGGCCAGGTGC
AGTGGCTCACGCTGTAATCCCAGCACTTTGGGAGGCCAAGGCAGGCGGATCACAAGGTC

13681 ACTACTCCCATTTTACTGATGGGAAACTGAGGCTCAGACATCATCTTCCCAGGCCAAAC
AGCTCTTCAATAGCAGAGCAGAGCTGTAAACCCACCTCTATAAGCCCTTTCACCCCCAC
CACACCATATGGAATTGGTTGCTAAACTGCTTCCTTGGGTACAGCAAATGGCATTGTGG
TTACAAGACCTTCCACGTGTGCTTCAAACAATGGGGTTTTGCCTAGACTAGTGCTTAGTA
GTAAGTGTATCACGGAAACACGGTCAGGACTCTTGGCGTCCATCTGATCGTGGGAGACCC
[T,G]

TCAGCATGAGCTGGATCCCCTCGGGGCTGTCTTTTCTTACATAAATGTTGCCTTTTGCC
CTTACTTGGTTTTTATTTTGTTCGCGACAATGGAAAACCTAATTTTTTTTTTATTA
AAGAAAAATCTATTCTGGCCAGGTGCAGTGGCTCACGCTGTAATCCCAGCACTTTGGGA
GGCCAAGGCAGGCGGATCACAAGGTGAGGATCGAGACCATCCTGGCTAACACAGTGAA
ACCCGCTCTACTAAAAATACAAAAAATAGCCGGGCGTGGTGGCGGGCGCTGTAGT

14336 CTTGCAGTGAGCCGAGATCACGCCACTGCACTCCAGCCTGGGCGACAAAGTGAGACTCTG
TCTCAAAAAAAAAAAAAAGAAAAATCTATTCTAAGTGAAGCAGTTTTTCCCAGTAGGTGG
CAGAACTAAATGCCATTATGCCATTTATAATTTAAGTGATTAAAGAGGAGTAGTATGTA
GTATATGCAAGGTCTAGCTCTAACAGCAGTGCAGTATAAATAGTAGAACTGACCTGATA

FIGURE 3P

TTACAGTATGAGAAACATGAAGGGGTTCTGTTTTGTGAGCTCTAAATTTATCTTCCATGT
[A,G]
TACTTCAAGGCTCTTCTCCCCAGTAGATTTTTATTTCATCTGAACTATAATTAGGTGGCCT
TTTTCCATTCTGAAAATAATTGGATCAAATGCATTTTAAAGTCCAGGGTCTGAAAGGTGG
AGGAATCCTTTCTCTTTACTGTTTCTAATTTAACTCCTTTTCATTTACTAGATTTTCAGT
CATGTCCAGAATTCATCTTTTCTAAAAGCTTTAATCTAGATTTAGAAATCTAAAATCTTT
TATTTATTTTTTTTTTCGTTGAAGTGCCCTGATTTTGTGGTGGTAAAGACTCCATTAGTA

14729 ATTTTAAAGTCCAGGGTCTGAAAGGTGGAGGAATCCTTTCTCTTACTGTTTCTAATTTA
AACTCCTTTTCATTTACTAGATTTTCAGTCATGTCCAGAATTCATCTTTTCTAAAAGCTTT
AATCTAGATTTAGAAATCTAAAATCTTTTATTTATTTTTTTTTTCGTTGAAGTGCCCTGAT
TTTGTGGTGGTAAAGACTCCATTAGTATCCACTTATACATTTCCCTGACTTTGCCTCTG
ACCAAACCTTACAGTATTCACATTGTAAGTGTGCAATAATAATAGCTAACATATTAATAC
[A,G]
CTGAATATTTGCTGTGTGCCTAAGCTAAGGATTTAATTCTCTTAAATCCTGTGAGGTAT
TTTATTTTACAGAAAAAGAACTGCTTAAAGAAAGTAACTTATCCAGGTCACACAAGTAA
CAATTGCAGAGCTGGAGTTTCAGATGAGGGCTGGCTTGCGCTGCCGTACAGAAAAGAGT
GCCCTAGAAATCGGTTCATTTGCATTTCCCGATTTTAGTTTAGCCAAATGAAAAATTCCT
TTTGGATTTATGAGTATAATCAGACAGTATACCTGTGAAATTAAAGTATTTGACTCTTTG

15124 GTAACCTATCCAGGTCACACAAGTAACAATTGCAGAGCTGGAGTTTCAGATGAGGGCTGG
CTTGCGCTGCCGTACAGAAAAGAGTGCCCTAGAAATCGGTTCATTTGCATTTCCCGATT
TTAGTTTAGCCAAATGAAAAATTCCTTTTGGATTTATGAGTATAATCAGACAGTATACCT
GTGAAATTTAAAGTATTTGACTCTTTGCTTGAAATAAGTAGGTTAAAAAGATTTGGGTGGC
CGGGCGCAGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGCAAGTAGATCA
[C,T]
TTGAGGTCAGGAGTTTCGAGACCAGCCTGACCAATATGGGGAAACCTCGTCTCTACTAAAA
ATACAAAAATTAGCCGGGCGTGGTGGTGCATGCCTGTAATACCAGCTACTTGGAGGCTGA
GGCAGGAGAAATCACTTGAAGCCAGGAGGCAGAGGTTACAGTGAGCTGAGATCACGCCACT
GCACTCCAGCCTGGGCAACAGAGCGGACTCTGTCTAACACAAAAAAGATTTGGGAAAA
CACTTTATTAATGAAGAGTTCCCTGACAAAGTGATTTTTTTGGGGAGAATTTTTATAATTG

15907 TTTTTAAAAATATTAAACATTAACTGCTCTTCTACCCACTCCAAGTCAAATAGCATTT
TTTCAGTCAGGTGTCTGGGAGCTCGATGCAAGATAACAAAATCTGGTCTCTGCCTCAGGG
AACATGAAATCTGTTTGGGGAAGCCAGAGCAAAAAATAAGGTTTTAATAGCAAGCTCTCA
CTAACTGCCCCTGAAATCCACCCACATCCTCCAGGAAGCCTTTCTTACCCCCAGTGC
CCTCAGGAGCTTCTCAAGGCAGGCCCTTCCAGAGCGCAGTGTGCTCCCCAGCTCACAG
[A,G]
AGATGCTCCCTACACGCTGCAGGAAAGTCCAGTGCCTGCAGCACAGGCTTCAGCAGCAGA
CTCGGGTTCTAGTCTCAGTCTGCTGATTCTAGTTGTGGAACCTGAGCAGGCGAAGTTAC
TAAACCTCTCTGTGCGTCAGCCTCCCAGGCTCGTTGCTTCAGGCCGAGTTAGGCTGTGT
GAACAGGAGAGTGGGGATGGGAACTAGGTATCTTAAAGCGGGGCAGAGTTTGGATGAGCG
GGCCACCCTTCGTATAGTTAGGAGGAAGATGACGGGAGGCATGGAAGCTGGGATAGCCAT

16341 GCGTCAGCCTCCCAGGCTCGTTGCTTCAGGCCGAGTTAGGCTGTGTGAACAGGAGAGTG
GGGATGGGAACTAGGTATCTTAAAGCGGGGCAGAGTTTGGATGAGCGGGCCACCCTTCGT

FIGURE 3Q

ATAGTTAGGAGGAAGATGACGGGAGGCATGGAAGCTGGGATAGCCATCCTGAGTCAGTGC
TAATTCTGACACTTCAGAACATCGAGTCAGTCTGACCTGCGAGTGAGCTTTTATTGACCA
CTTAGAAACTATTAGCACCTTGGACAACTACTTTCTTTGAGACCTGGTTGCTTCATGTC
[-,G,T]

GCGATGGGAAAACTGATACTTAACTTGCAGATAGTGGTGAATCAAAAGTAGTATATGTGA
AGTACTCACACTGCGGAGCATTAGCCATCGTCCCATCCTACTTCTACCTTTTACATA
TTGTAATATGAAAGCTAAACCATTTCTCGATGTGAGTCAGTTTTAATCGGCTACATAGTG
AGTGGCATTGATTTTTAAAAATGTCAACTTGGGATCTGTACCATGCTACTTACCATTGT
TATGTCACACTGTTGAATGTCGGACCTGGTTTGTCTTCTCCAGATGGTATGTTACTTA

16786 TCTCGATGTGAGTCAGTTTTAATCGGCTACATAGTGAGTGGCATTGATTTTTAAAAATGT
CAACTTGGGATCTGTACCATGCTACTTACCATTTGTATGTCACACTGTTTGAATGTCGG
ACCTGGTTTGTCTTCTCCAGATGGTATGTTACTTACGGTCATGAACTGATTTGGAAGAA
CAGGGAGCCGCTAGTGAAAATCTGGCATGAAATAAGGACTAATGGCCCCAAAAAGGAGG
TGGCTCTAAGTAAACTGGGATTGGACAGTAGTGGTGCATCTGGTCCTTGCCGCTGAGA
[G,C]

CCCCAGGAGACATCGGCTAGAGTGACCATGGCTATGCTCCCGTCTGGAAGATGCCAGCAT
CTGGCCTCCCACTGTTTTAGCTGTGTCCCCAGTCCGTGTCTTTTAGAATGTGAATGA
TGATAAAGTTGTGAAATAAAGTTTCTATCTAGTTTGTAAAGCAGATGTGTGTCTCTC
TTTAAGGGGCCGACACGGCTCTGGCATTGTGCTTTGGTTGTGCAATTGACAGGACCTGGG
GAGAGTGACCCCTGAAAGGCCTGATCAGAACATGAAGGCGCTGGTTGCCTGTCTTTGGAC

17159 TGTTTTAGCTGTGTCCCCAGTCCGTGTCTTTTAGAATGTGAATGATGATAAAGTTGT
GAAATAAAGTTTCTATCTAGTTTGTAAAGCAGATGTGTGTGTCTCTTTAAGGGGCCG
ACACGGCTCTGGCATTGTGCTTTGGTTGTGCAATTGACAGGACCTGGGGAGAGTGACCCC
TGAAAGGCCTGATCAGAACATGAAGGCGCTGGTTGCCTGTCTTTGGACCCTCCAGTGCCT
CTGCTTAGCCTTCACTCTTCTTGCCTCCCCCTCCCCCTGGGTTGGCTGCACATAAAAGTC
[G,A]

AGAGTATCCCCTCTCCAGCACAACTGAAATAACAGCTGCAGTATTTTCTCAATTTTCTAG
GAAAGGTAGTGTCTTCTGGCAGTGAGTGGCATATACAAAAAGCTATTTTCTAGTTTGTCT
TTCTAGGTTCAATTTGTAGATAAATTAAGAGGTAGAAAGAAGTGATTTGGGTAAATTCAG
ACTTGAAATCTGAGCCGAATTTATCTTCTGTTTGAAGTGTTCTAATTGAAGCGTCTCA
CTGAAATAGCAGATAGTGGCTGTCGTCGTACAGCCCTCACTGTTGTGGAATTCATGTT

17976 AAAAGGAGTGGGAGTGGTGCCACCTACCAGGCAAGTGAGAACTGCATGGCAGCACGCG
CCCAGCACATAGAAATTGTCCAGTATTTGGCAGTCCCTCATATCCTTCTTCCATCAGGCT
GGACTTGTCTTCTACTATGATTTACAGTTATTCTTCCCAGGCACAGGATTCTGTTCTAAAC
TCGTATCACTTCTAGGGGAGAGAGTTATCTTAGCCATCATTTTGCCAGCGAGGAAACGGC
ACACGTGGTGTAGGGGCACTGCCAAGGTCACAATGCTTTGCTCTGACATCTGCTAACAA
[-,T,C]

TGCAACACAGATGAGGCAAGATGCGTTTTCCAGAGATGGGATAGGAGGCTGAGTTCATAG
GGACATTCCCTCTAGAGCCCAACATTAATTCACATCGTGCTTTGGGCAGACCAGGCAAAG
AGGCAATGAAGACATCTCTGTGTCCCTGCTTTGTGACTGGGAAAAAGTTAGAAGTCCCTG
TAGCATCTCCTGGTCCCTAAAACCCCTCAATGCTGGAGCCTCTGTGCATGGCCTGGGGAG
GCCAGAACCTGGCTGTGGCCGGAGAAGCCTTGCTGTCCACAGCTCCCTCCTGATTGCCCA

FIGURE 3R

- 18001 TCACCAGGCAAGTGAGAACTGCATGGCAGCACGCGCCAGCACATAGAAATTGTCCAGTA
TTTGGCAGTCCTTCATATCCTTCTTCCATCAGGCTGGACTTGTTTCTACTATGATTTACA
GTTATTTCTTCCCAGGCACAGGATTCTGTTCTAAACTCGTATCACTTCTAGGGGAGAGAGT
TATCTTAGCCATCATTTTGCCAGCGAGGAAACGGCACACGTGGTGTAGGGGCACTGCCCCA
AGGTCACAATGCTTTGCTCTGACATCTGCTAACAACGCAACACAGATGAGGCAAGATGC
[G,A]
TTTTCCAGAGATGGGATAGGAGGCTGAGTTCATAGGGACATTCCTCTAGAGCCCAACAT
TAATTCACATCGTGCTTTGGGCAGACCAGGCAAAGAGGCAATGAAGACATCTCTGTGTCC
CTGCTTTGTGACTGGGAAAAAGTTAGAAGTCCCTGTAGCATCTCCTGGTCCCTAAACCC
CTCAATGCTGGAGCCTCTGTGCATGGCCTGGGAGGCCAGAACCTGGCTGTGGCCGAGA
AGCCTTGCTGTCCACAGCTCCCTCCTGATTGCCACAGAGGGTGCTTCACTTCTCCTCTT
- 18021 GCATGGCAGCACGCGCCAGCACATAGAAATTGTCCAGTATTTGGCAGTCCTTCATATCC
TTCTTCCATCAGGCTGGACTTGTTTCTACTATGATTTACAGTTATTCTTCCCAGGCACAG
GATTCTGTTCTAAACTCGTATCACTTCTAGGGGAGAGAGTTATCTTAGCCATCATTTTGC
CAGCGAGGAAACGGCACACGTGGTGTAGGGGCACTGCCCCAAGGTCACAATGCTTTGCTCT
GACATCTGCTAACAACGCAACACAGATGAGGCAAGATGCGTTTTCCAGAGATGGGATAG
[G,T]
AGGCTGAGTTCATAGGGACATTCCTCTAGAGCCCAACATTAATTCACATCGTGCTTTGG
GCAGACCAGGCAAAGAGGCAATGAAGACATCTCTGTGTCCCTGCTTTGTGACTGGGAAAA
AGTTAGAAGTCCCTGTAGCATCTCCTGGTCCCTAAACCCCTCAATGCTGGAGCCTCTGT
GCATGGCCTGGGAGGCCAGAACCTGGCTGTGGCCGAGAAGCCTTGCTGTCCACAGCTC
CCTCCTGATTGCCACAGAGGGTGCTTCACTTCTCCTCTTGGCTTCTCTGGGGACCCGCG
- 18022 CATGGCAGCACGCGCCAGCACATAGAAATTGTCCAGTATTTGGCAGTCCTTCATATCCT
TCTTCCATCAGGCTGGACTTGTTTCTACTATGATTTACAGTTATTCTTCCCAGGCACAGG
ATTCTGTTCTAAACTCGTATCACTTCTAGGGGAGAGAGTTATCTTAGCCATCATTTTGCC
AGCGAGGAAACGGCACACGTGGTGTAGGGGCACTGCCCCAAGGTCACAATGCTTTGCTCTG
ACATCTGCTAACAACGCAACACAGATGAGGCAAGATGCGTTTTCCAGAGATGGGATAGG
[A,G]
GGCTGAGTTCATAGGGACATTCCTCTAGAGCCCAACATTAATTCACATCGTGCTTTGGG
CAGACCAGGCAAAGAGGCAATGAAGACATCTCTGTGTCCCTGCTTTGTGACTGGGAAAA
GTTAGAAGTCCCTGTAGCATCTCCTGGTCCCTAAACCCCTCAATGCTGGAGCCTCTGTG
CATGGCCTGGGAGGCCAGAACCTGGCTGTGGCCGAGAAGCCTTGCTGTCCACAGCTCC
CTCCTGATTGCCACAGAGGGTGCTTCACTTCTCCTCTTGGCTTCTCTGGGGACCCGCGA
- 18042 ACATAGAAATTGTCCAGTATTTGGCAGTCCTTCATATCCTTCTTCCATCAGGCTGGACTT
GTTTCTACTATGATTTACAGTTATTCTTCCCAGGCACAGGATTCTGTTCTAAACTCGTAT
CACTTCTAGGGGAGAGAGTTATCTTAGCCATCATTTTGGCAGCGAGGAAACGGCACACGT
GGTGTAGGGGCACTGCCCCAAGGTCACAATGCTTTGCTCTGACATCTGCTAACAACGCAA
CACAGATGAGGCAAGATGCGTTTTCCAGAGATGGGATAGGAGGCTGAGTTCATAGGGACA
[T,G]
TCCCTCTAGAGCCCAACATTAATTCACATCGTGCTTTGGGCAGACCAGGCAAAGAGGCAA
TGAAGACATCTCTGTGTCCCTGCTTTGTGACTGGGAAAAAGTTAGAAGTCCCTGTAGCAT
CTCCTGGTCCCTAAACCCCTCAATGCTGGAGCCTCTGTGCATGGCCTGGGAGGCCAGA
ACCTGGCTGTGGCCGAGAAGCCTTGCTGTCCACAGCTCCCTCCTGATTGCCACAGAGGG

FIGURE 3S

TGCTTCACTTTCTCCTCTTGGCTTCTCTGGGGACCCGCGATCACTGCCTTCAAGGCCATG

18375

GCTTTGGGCAGACCAGGCAAAGAGGCAATGAAGACATCTCTGTGTCCCTGCTTTGTGACT
GGGAAAAAGTTAGAACTCCCTGTAGCATCTCTGGTCCCTAAAACCCCTCAATGCTGGAG
CCTCTGTGCATGGCCTGGGGAGGCCAGAACCTGGCTGTGGCCGGAGAAGCCTTGCTGTCC
ACAGCTCCCTCCTGATTGCCACGAGGGTGCTTCACTTTCTCCTCTTGGCTTCTCTGGGG
ACCCGCGATCACTGCCTTCAAGGCCATGCACTCCCTGGCCCGTGGGCCTCTTGGGCTGTG
[C,T]
CGCCTCCACTGGCATCTGAAGTGTGGGGTACCTAGGAACATGCCGTGGCTGCCGTCTCCC
TCATTCCATACACTTCTTGAGTGGGTGCACTTGCTGAAGCCTCAGTTATCTGTGAGGATT
CTGAGCTCCAGACCCACAGAATCTCTCTGTACTCTTAGTAAATGTGTCTACTGCAACACA
CGCATGGTTCAGGCTCTGGGACCACCCCCCGCCCTGCACAGGCCCTCAAATAGCACT
CGGCTTAAGGAGTGACACGAGCAATCGGTGAAGTCTGAAACCCGGAGCCATTGAGATCT

19244

CTAGATGGTCACTACACTCAGGGAGTTGGGGATGGCTCAGAGCTGTTAACAGAGAGGGGA
CTGCCCAGGAGGACCTGCGTGAGGGGTGGGGTGGGATGACAAGGAACCAGCTCTGGGAG
TTGAAAGACCTGGATTCAAGTCTCAACCCAAGCCCTGGCCAGCTCTGGGACCCCGGACAA
GTCGGCCTCACTCTCTGCCCCCTCAGTGGGCTCCTGTGTAGATGGGGATAATGATGGCTTT
ATATCCTGAGAATGTGGGGAGGGGATTAAGTGGCCAAAATACCTGAGAGTGCGCACTCAG
[T,C]
GCCTGGCTCAGCAAATGCCCTTGTTCCCTCCTTCCCTCTCCCCAGAACCCCTCCTCCCCT
TCTTCTTCTTTTTTTTTTTTTTTTTTTTGACCCAGAGTCTTGCTATGTTGCCCAGGCTGGA
GTGCAGTGGCACAATCTCGGCTCACTGCAACCTCCACCTCCTGGCTTCAGGCAATTCTTG
TGCCTCAGCCTCTCGAGTAGCTGGGATTACAGGCAGGCACCATCACGCCCGGCTAATTTT
TTTTTTTTTTTTTTGTAGTAGAAATGGGATTTACCATATTGGCAGGATGTTCTCGATCT

Chromosome map:

Chromosome 3